USING OF RAW RADIOSONDE DATA FOR QC/QA

A. Kats, V. Grinchenko, A. Balagurov

FGUP KOMET, Roshydromet
3, build.6, Pervomaiskaya Street, Dolgoprudny, 141700, Russian Federation
Tel.: +(7 495) 408 6104, Fax:+(7 495) 408 6865
E-mail: komet.kats@mtu-net.ru

ABSTRACT

Presented are examples and discussed are area of raw data usage for quality control/assurance of upper-air ground equipment, radiosondes and data processing software:

- radiosonde telemetry consistency;
- radiosonde position tracking consistency (especially important for systems with upper-air radar);
- specific radiosonde-dependent issues (MRZ transducer, RS90/RS92 dual HUMICAP sensors);
- identifying unreasonable QC (strong inversions, supersaturation over ice);
- identifying data gaps and inappropriate interpolation.

It is stressed the importance of archiving and analysis of raw data as well as necessity of good cooperation between manufacturers, oversight scientists and operational stuff to establish relevant procedures.

Introduction

Modern automated upper-air processing often deeply hides details of measuring and processing. This is a great advantage unless hidden are sporadic (operators mistakes, equipment malfunction and failures, external interference), occasional (improper maintenance, production variability) and/or inherent (design, algorithms) problems. Some of them are:

– inadequate surface observation and improper sensors conditioning/exposure before and just after release of a balloon;
– radiosonde and tracking system malfunction;
– radiosonde telemetry reception and tracking position inconsistency;
– loosing valid information due to QC errors of the second kind (rejecting plausible values because of unnecessarily rigorous control);
– introducing bogus data (improper interpolation of signal gaps).

Fortunately, now automated systems provide (or can or could provide) lot of helpful information that allows to open and look inside their “clever black box”. As such information usually never reaches end users it is usually rather informally referred to as “raw” data.

“Raw”data – what they are

In the broad sense “raw” data are recognized hereinafter as the information used at all steps of data reduction /WMO 2008, Chapter III.3/ from radiosonde sensor and tracking system readings preceded to eventual profiles of meteorological values reported to end users in agreed formats. In the terms of /WMO 1992, 2007/ they include proper raw data and primary data (i.e. Level 0 and Level I data levels) but the boundary between these two categories are quite relative.

Proper raw data are first and foremost attributed to absolute or elapsed ascent time:

– radiosonde telemetry information: raw sensors and internal references signals (in the terms of capacitance, resistance, voltage, frequency etc), battery voltage, failures rate, carrier frequency, signal-to-noise ratio, hardware status (e.g. RS92 HUMICAP heating), etc;

– radiosonde position tracking: signals and following angular error of radar tracking system, navaid signals (including e.g. amount of available transmitting stations or satellites in view), etc.

Primary data are:

– uncorrected readings of surface instruments;

– raw PTU measurements after applying the calibration functions to telemetry readings (including auxiliary information such as e.g. internal radiosonde or humidity sensor temperature) and radiosonde position with native sampling rate[1], often determined by radiosonde telemetry cycle duration;

– unedited (unfiltered) and uncorrected derived quantities (such as wind);

– components of corrections and auxiliary quantities necessary for correction (e.g. vertical ascent rate and solar elevation), quality control flags, information related to operator intervention (e.g. significant levels editing).

Under curtain circumstances even meteorological parameters[2], corrected, edited (filtered) and interpolated to a regular time grid, could be considered as “raw” data unless they are transmitted[3] with help of reports using TDCF.

For convenience as “raw” data in abovementioned sense could be referred to internally available metadata most of which will be available for end users of upper-air data in future in TDCF-coded messages. They include:

– location of station (to which barometric pressure reports at the station refer) and location of radiosonde release point, magnetic declination;

– tracking radar baseline distance and angular corrections and elevation of aerial above the station height;

– standard surface sensors corrections (e.g. for mercury barometer or wet and dry bulb thermometer);

– ground system hardware and software version;

– balloon type and free lift;

– radiosonde serial number, calibration coefficients and ground-check results;

– and all like that.

All this information is worth to be stored along with each radiosonde ascent data archive as storage capacity of modern computers does not put any obstacles to that.

It’s rather difficult to guess a priori which kind of information can be helpful in a particular situation while content and amount of data available are highly system dependent. Therefore it may have sense to encourage manufacturers to provide opportunity to record all possible information as well as to document them properly to avoid forcing users to dig them out.

---

[1] Reported time grid may be different from one of edited data

[2] Which are already Level II data in the terms of /WMO 485, WMO 488/.

[3] In the most of modern systems such high-resolution data are recorded for off-line collection and archiving in data centers, however this is not the case always.

Examples

Shown below some practical examples of “raw” data analysis are only a part of that could be imagined although hopefully sufficient to demonstrate its usefulness. The idea was not to concentrate on particular details but instead to present rather a wide panorama. In most cases for “raw” data analysis it was used specially developed dedicated software. Initially it was made for AVK-ARM Aerolog data processing system /WMO 2000/ but now is available for another automated upper-air systems, used in Russia.

Telemetry raw data

Figure 1.a demonstrates a gap (absence of telemetry periods) in radiosonde MRZ-3 humidity channel. A gap most likely was caused by a poor contact in rheostat that converts goldbeater skin humidity sensor tension into resistance (as shown on Figure 1.b loss of contact also influenced a duration of telemetry cycles which is controlled by a transducer multiplexor). As such gaps happened at the station several times in radiosondes from the same production batch the most likely reason was a production flaw. Such records may give a station a ground to raise a claim to manufacturer.

Another example of a problem in electric circuit is demonstrated on Figure 2 which shows telemetry periods of radiosonde RF95 (with Vaisala RS80 sensor unit). Investigation of such records allowed identifying and reproducing in laboratory the reason of the problem – a break in temperature sensor circuit caused by mechanical damage of a sensor boom. The problem was removed owing to modification of the sensor boom mounting support and transportation package design.

Figure 3 demonstrates a problem of radiosonde MRZ-3 transducer which was identified using raw data. Radiosonde MRZ-3 has 2 telemetry channels for the same temperature sensor. Due to production problem a big batch of radiosondes had an electronic problem in transducer resulting in erroneous reading in one of the channels (an extreme case is shown). A gap in data in this flight was caused by a reception problem.

In an example shown on Figure 4 data processing software did not cope with proper identification of radiosonde MRZ-3 telemetry channels in situation with poor reception – spurious values were reported to the next level of data processing instead of indicating missing data.

Radar raw data

Not modernized with solid-state microwave modules AVK radar used on Russian upper-air network has two transmitters with low and high power. System (or operator) switches from low to high power transmitter at a distance about 3 km. When radar is not properly adjusted a discontinuity in measured slant ranges may take place at this time. Before introducing PC-based data processing systems, allowing raw data recording, it was quite difficult to identify a problem. Figure 5 shows a presence of the problem quite obviously.

Another typical problem of AVK radar maintenance is necessity of proper adjustment of angular measurement. Actually they are made with two transducers: rough and fine selsyns (like hour and minute hands in watches[4]). Improper adjustment results in error equal to the whole scale of fine selsyn. Figures 6.a and 6.b apparently demonstrate cases where respective maintenance is necessary.

Radiosonde tracking in boundary layer, when just after release angular velocity of radiosonde movement relative to aerial may be quite high (when a radiosonde approaches a station) is a weak point of radar, especially if radar is not maintained in a proper way. Analyzing of tracking data

---

[4] In case of watches minute dial is equal to 1/24 of hour dial, in case of AVK radar the whole scale of fine readings is equal to 1/32 of the whole range of 360°, i.e. 11.25°.

helps to monitor radar performance and to train operators locating a radiosonde after a loss (see Figure 7).

Quite an obvious and intuitive indication of radar performance is instantaneous height, calculated from raw elevation and slant range. Even taking into account natural variability of vertical ascent rate of a balloon, balloon height is rather a smooth and consistent function of time (see Figure 8). Therefore, rough errors in slant range or elevation tracking are immediately visible (see Figure 9.a), while an experienced person can evaluate even less evident cases[5]. To facilitate such an assessment in evaluation software it was introduced calculation of vertical ascent rate Va and auxiliary quantity dZ, corresponding to deviation of instantaneous height from height, calculated in assumption of constant ascent rate.

Vertical ascent range is estimated[6] roughly as

$$Va(t)=[Z(t)-Z(t-300)]/300$$

where

t – elapsed ascent time [s],

Z(t) – instantaneous height.

Auxiliary quantity dZ is calculated as

$$dZ(t)=Z(t)-Va_{mean}\cdot t$$

where

Va_{mean} – an estimate of mean vertical ascent rate

$$Va_{mean}=Z_{max}/t_{max},$$

where

Z_{max} – maximal height,

t_{max} – time of reaching maximum height.

Respective estimates for a problematic ascent are shown on Figure 9.b while on Figures. 10.a and 10.b for comparison are shown same quantities calculated from raw GPS heights and heights, derived from raw PTU data, of Vaisala RS92 radiosonde. Two more examples are shown on Figures 11 and 12.

There are also useful auxiliary parameters, recorded by ARM Aerolog, such as angular tracking error (see Figure 13) and flag, indicated that radiosonde signal-to-noise ratio is below the nominal limit. The former is useful to monitor performance of tracking automatics and, in case of necessity, to locate a position of unwanted radio-frequency interference source. The latter is useful to monitor both sensitivity of radar receiving system and signal power of radiosondes.

In many cases raw data were also found to be more informative for comparison of upper-air radars rather then comparison of processed wind results (see Figure 14).

Primary data

Surface observation shall be made as close in time to radiosonde release as possible, radiosonde sensors shall be ventilated. Figure 15 demonstrates an ascent where this might be not the case.

Rather a specific problem of automated data-processing system is loosing valid information when quality control subsystem rejects plausible or possibly plausible values. Such cases should be

[5] Horizontal trajectory projections evaluation sometimes also could be useful although it's not so evident as height.

[6] Va estimate is obviously not perfect, but for quality assessment it's not essential.

carefully examined and appropriate software modifications undertaken. An example of the QC second order error is shown on Figure 16.

Another similar case is quality control and representation of relative humidity where preservation of raw and primary data may have critical value for climatology. Not only super saturation over water is truncated (which is predetermined[7] by conventional way of reporting humidity in upper-air messages as positively defined dew-point depression in accordance with Code table 0777), but often humidity well above super-saturation is also truncated (see Figure 17). Truncated information is lost for the further analysis and possible correction thus influencing climatological results for water vapor quantities distribution. Less apparent case is truncating extremely zero or even negative humidity values to 1% RH, that prevent any subsequent corrections that still might yield plausible results. Last but not least point regarding relative humidity is reporting resolution – 0.1%RH is much more favorable for application of time-lag correction /WMO 2005/.

An opposite case is when quality control system fills in data gaps that should be carefully examined during acceptance of new systems or software modifications because limits used for interpolation could be an issue. Of course, sensitivity and robustness to outliers is even more important. Examples of ascents demonstrating how essential can be these issues are shown on Figures 18 and 19. Analyses of such cases should result in accurate specification of acceptable amount of data loss. Interpolated gaps must be reflected by respective quality/status flags when TDCF are used.

RS90/RS92 alternatively heated H-HUMICAP humidity sensors

One rather special but practically interesting case of primary data use is an analysis of Vaisala RS90/RS92 humidity measurements, equipped with two H-HUMICAP heated sensors /Paukkunen, 1995/. Using raw and primary data allow, for example, studying consistency of measurements between both sensors at different conditions (see Figure 20).

Lessons

Manual "raw" data analysis was proved to be very useful to reveal problems of operational sounding. However, its application for operational quality control is rather limited due to the need of timely onward transmission on the GTS, although sometimes it may be useful to undertake appropriate reprocessing and re-send a corrected message to GTS or even undertake the second release. Not less important it requires relevant expertise and experience. Therefore, its scope is mainly non real-time maintenance support and performance monitoring by both site technical managers and supervising authorities. Auxiliary software development was found to be very helpful to facilitate "raw" data analysis for quality assurance.

Only "raw" data provide possibility of adequate re-processing if data entry, calibration, correction or software problems were identified and fixed. The lower the level of raw data archived – the more fundamental and adequate retrospective corrections could be done. For example, Level I data are enough if mistake was made in entering surface pressure while proper raw data are mandatory when error was made in entering of calibration data. In most cases, such reprocessing is actual at preparation upper-air data to climatological processing (e.g. producing CLIMAT TEMP messages), especially on GUAN and on established GRUAN sites.

Special case is use of "raw" data analysis for test of new equipment and software. "Raw" data are irreplaceable when detailed analysis of a new system performance is required.

Last but not least a good cooperation between manufacturers, operational stuff and oversight scientists is required to establish and implement relevant effective procedures

[7] Modern TDCF codes concept doesn't prevent reporting >100%RH but it seems currently implemented BUFR upper-air messages inherit this historic practice.

Acknowledgments

Data files, used for this paper were provided by courtesy of numerous Russian upper-air stations and DWD Lindenberg Meteorological Observatory - Richard Aßmann Observatory.

References:

World Meteorological Organization, 1992: Manual on the Global Data-processing System. WMO-No. 485, Geneva.

Paukkunen, A., 1995: Sensor heating to enhance reliability of radiosonde humidity measurement. Proc. 11th Conf. on Interactive Information and Processing Systems for Meteorology,Oceanography, and Hydrology, Dallas, TX, Amer. Meteor. Soc., 103–106.

World Meteorological Organization, 2000: A. Balagurov, V. Grinchenko, A. Kats. New Instruments for Russian Upper-air Network. Papers Presented at the WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (TECO-2000). Beijing, China, 23-27 October 2000. WMO, Instruments and Obs. Methods Report No. 74.

World Meteorological Organization, 2005: A. Kats, U. Leiterer, H. Dier. Time-lag Correction of Operational RS80-A Radiosonde Humidity. Papers Presented at the WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (TECO-2005). Bucharest, Romania, 4-7 May 2005. WMO, Instruments and Obs. Methods Report No. 82.

World Meteorological Organization, 2007: Guide on the Global Observing System. WMO-No. 488, Geneva.

World Meteorological Organization, 2008: Guide to meteorological instruments and methods of observation. WMO-No. 8, Geneva.

Figures[8]

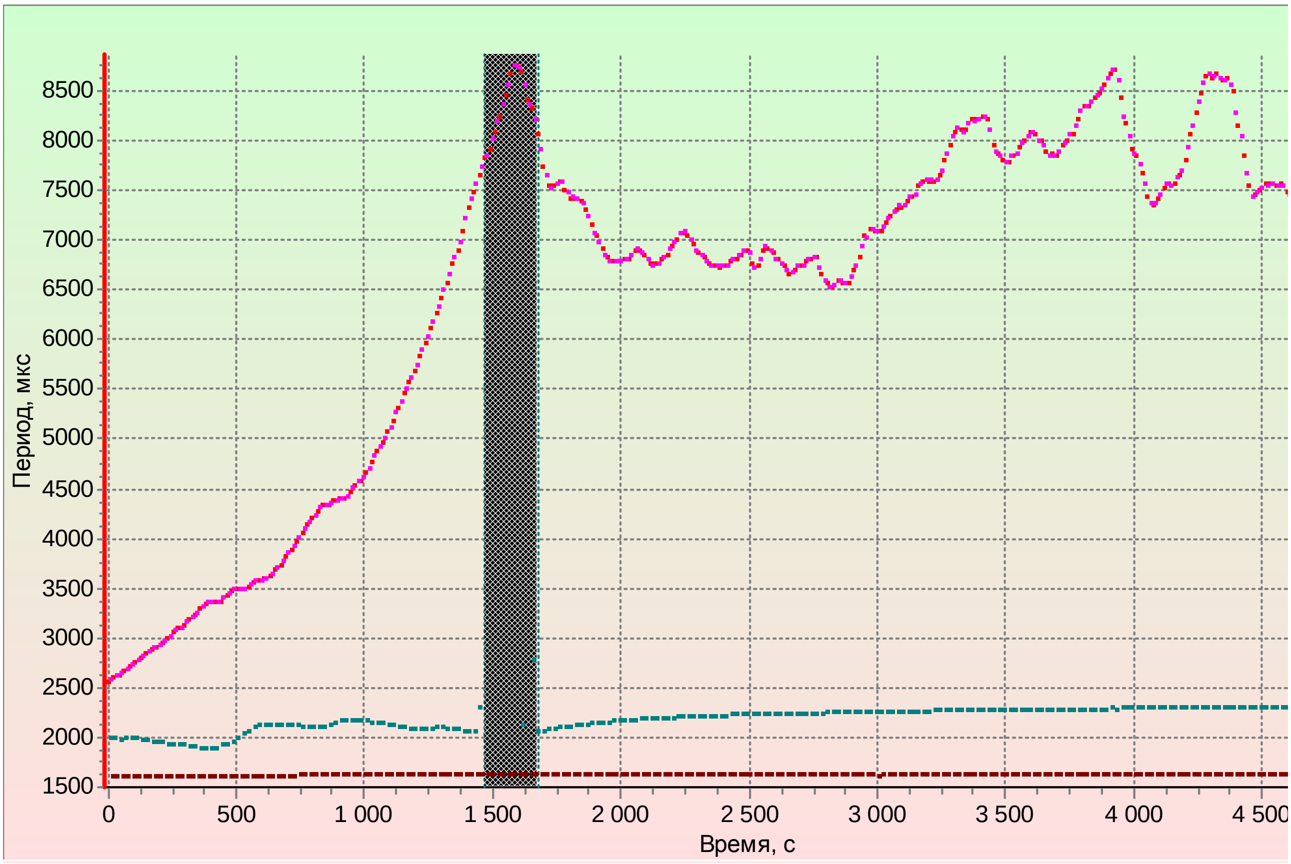

Figure 1.a. Telemetry periods of radiosonde MRZ-3 [mks] (brown – reference channel, pink and red – temperature channels, cyan – humidity channel) – a gap in humidity channel.

[8] In most of figures abscissa is elapsed ascent time in seconds.

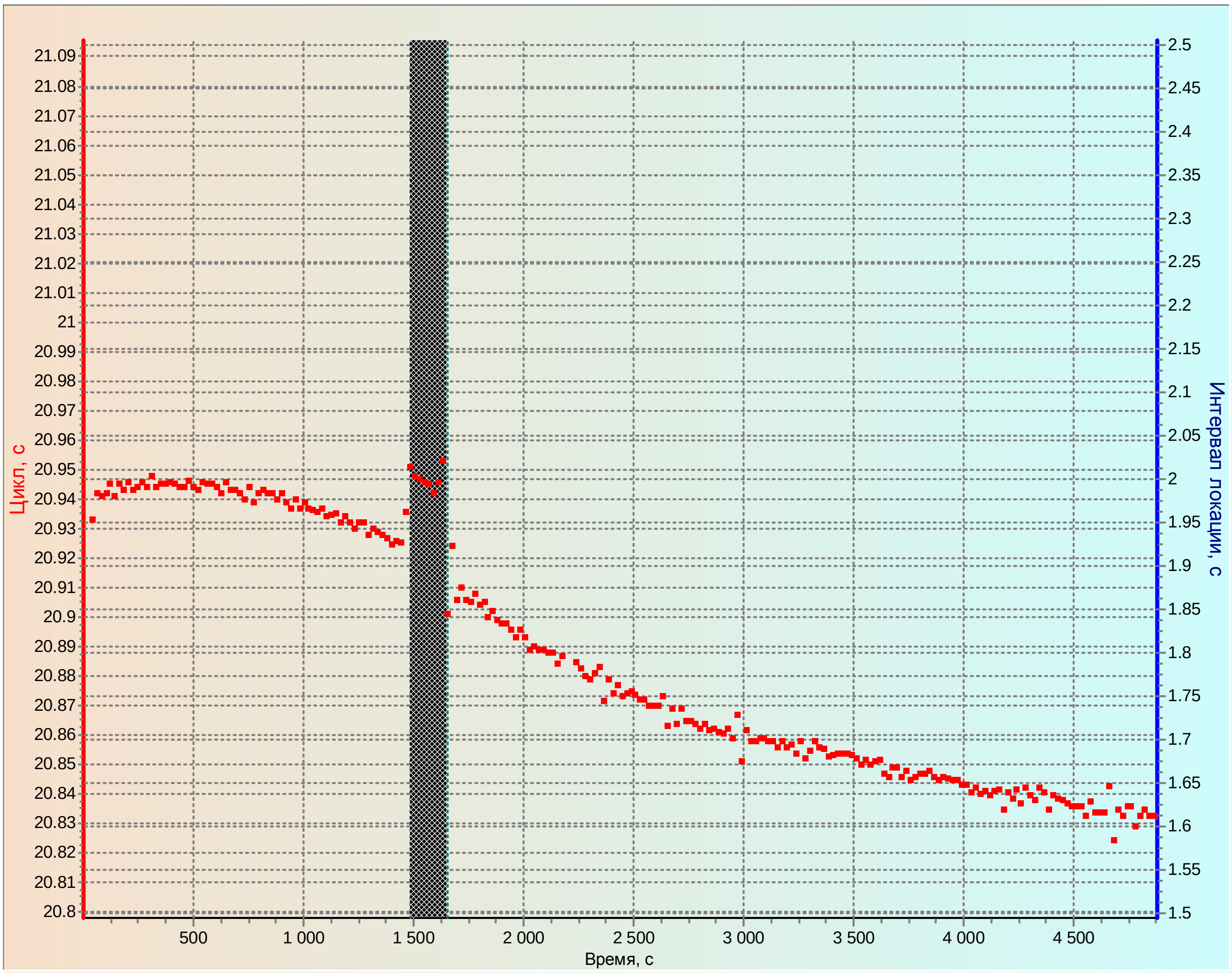

Figure 1.b. Same ascent as on Figure 1.a. Duration of radiosonde MRZ telemetry cycle [s] (red).

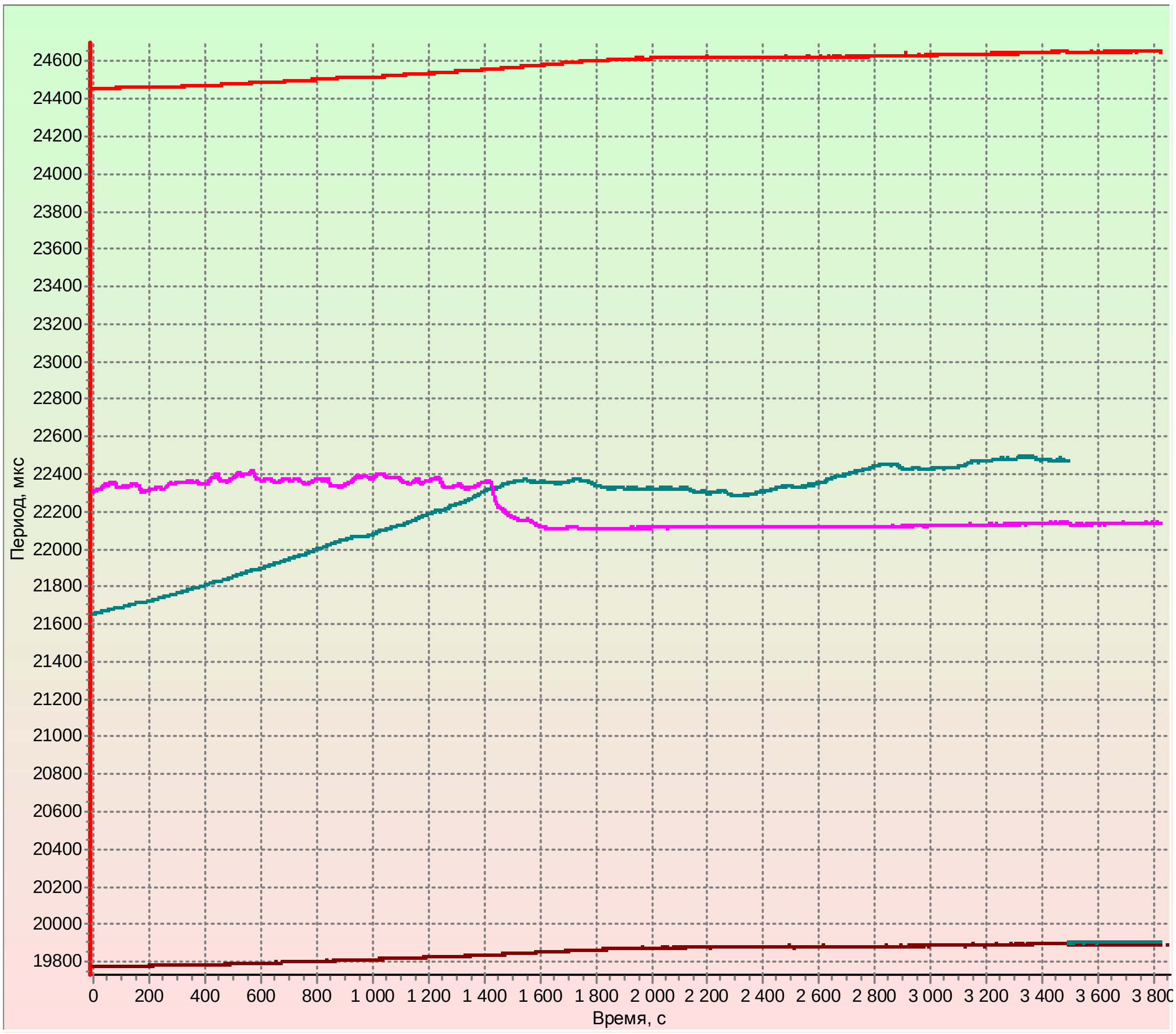

Figure 2. Telemetry periods of radiosonde RF95 [mks*3] (brown and red – reference channels, cyan – temperature channel, pink – humidity channel) – break in temperature sensor circuit.

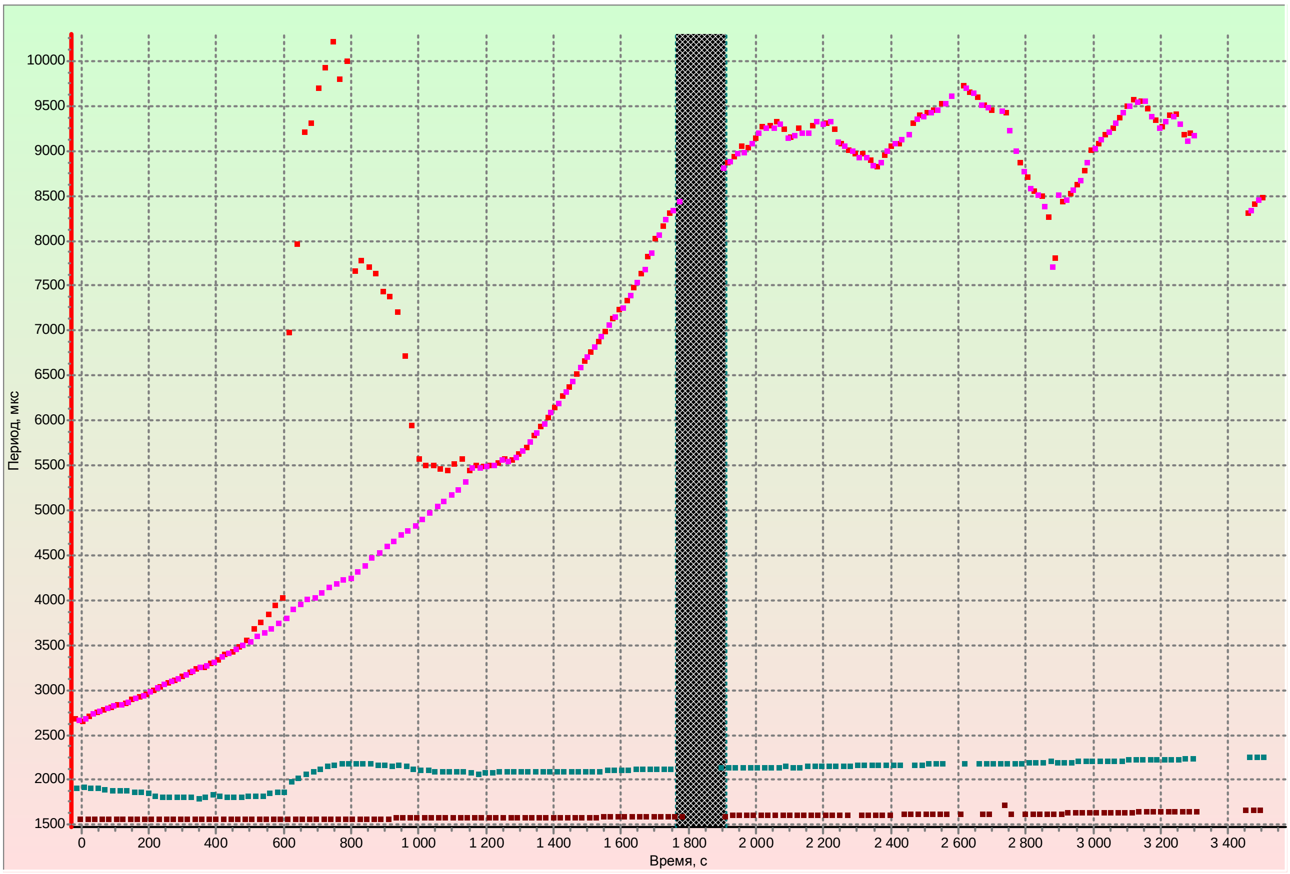

Figure 3. Telemetry periods of radiosonde MRZ-3 [mks] (brown – reference channel, pink and red – temperature channel, cyan – humidity channel) – transducer malfunction.

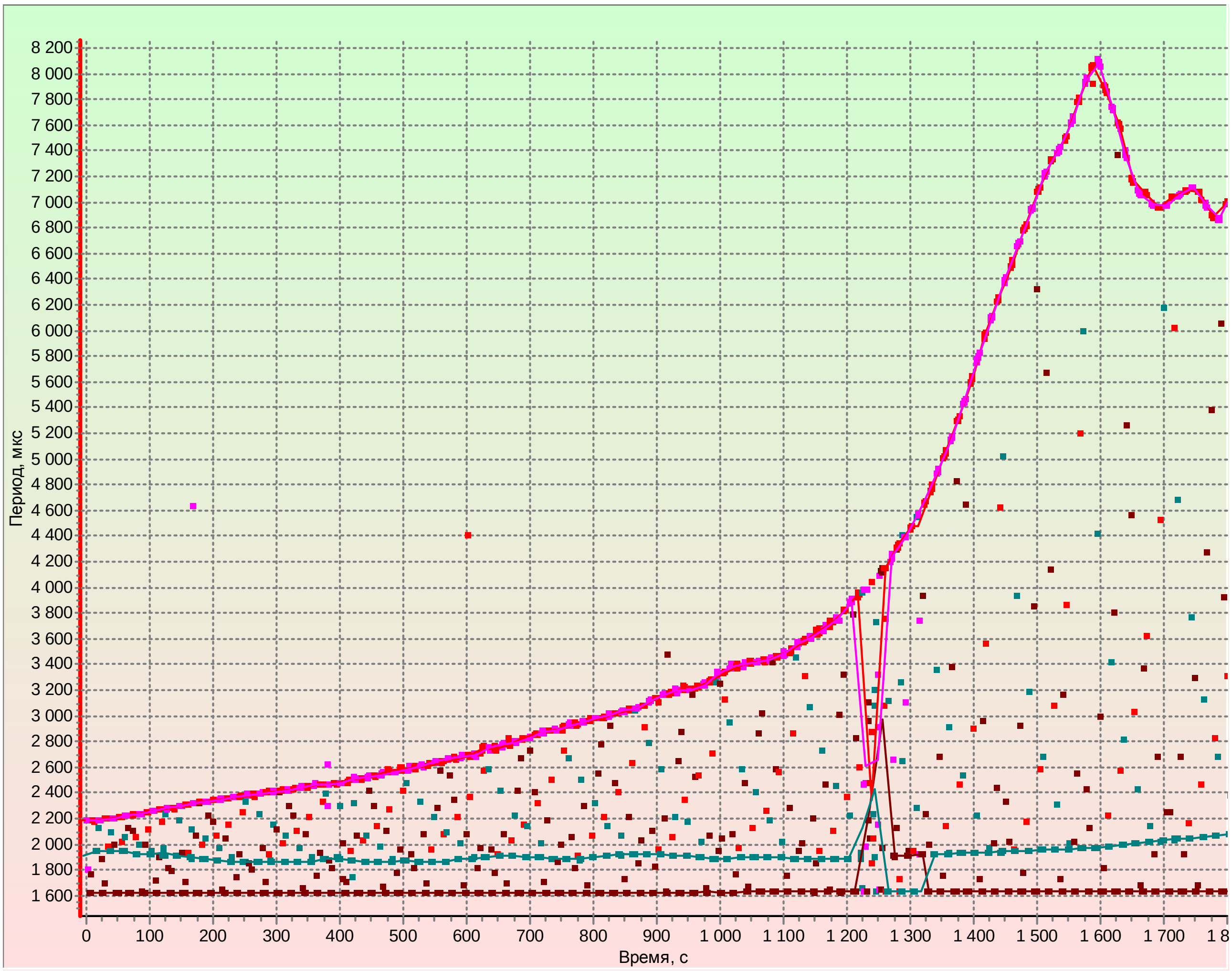

Figure 4. Telemetry periods of radiosonde MRZ-3 [mks] (brown – reference channel, pink and red – temperature channels, cyan – humidity channel; points – one second averages; lines – channel averages) – a software problem with channel identification.

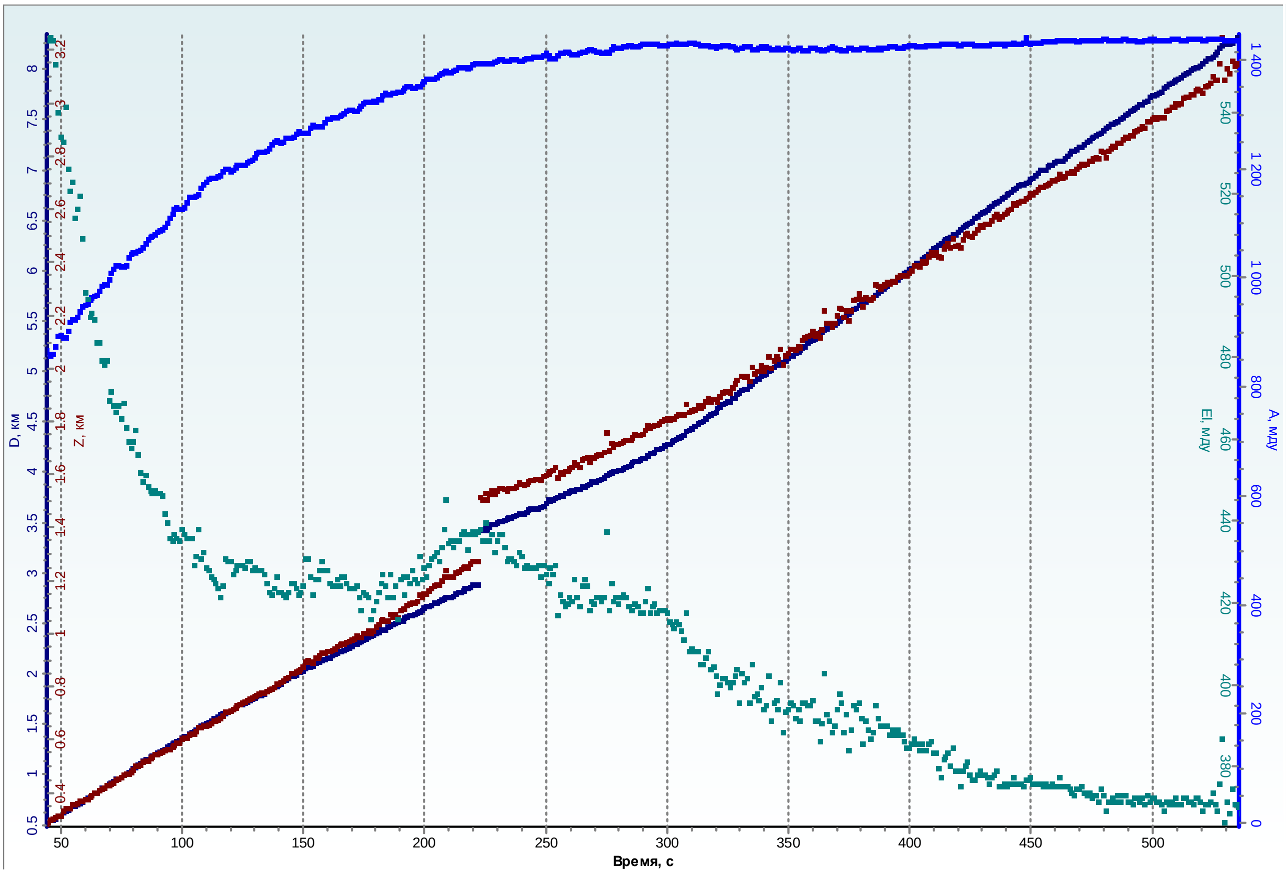

Figure 5. AVK radar tracking data (dark blue – slant range [km], brown – instantaneous height [km], blue – azimuth [1x0.06°] and cyan – elevation [1x0.06°]) – a discontinuity in slant range due to switching from low to high power transmitter after 3 km.

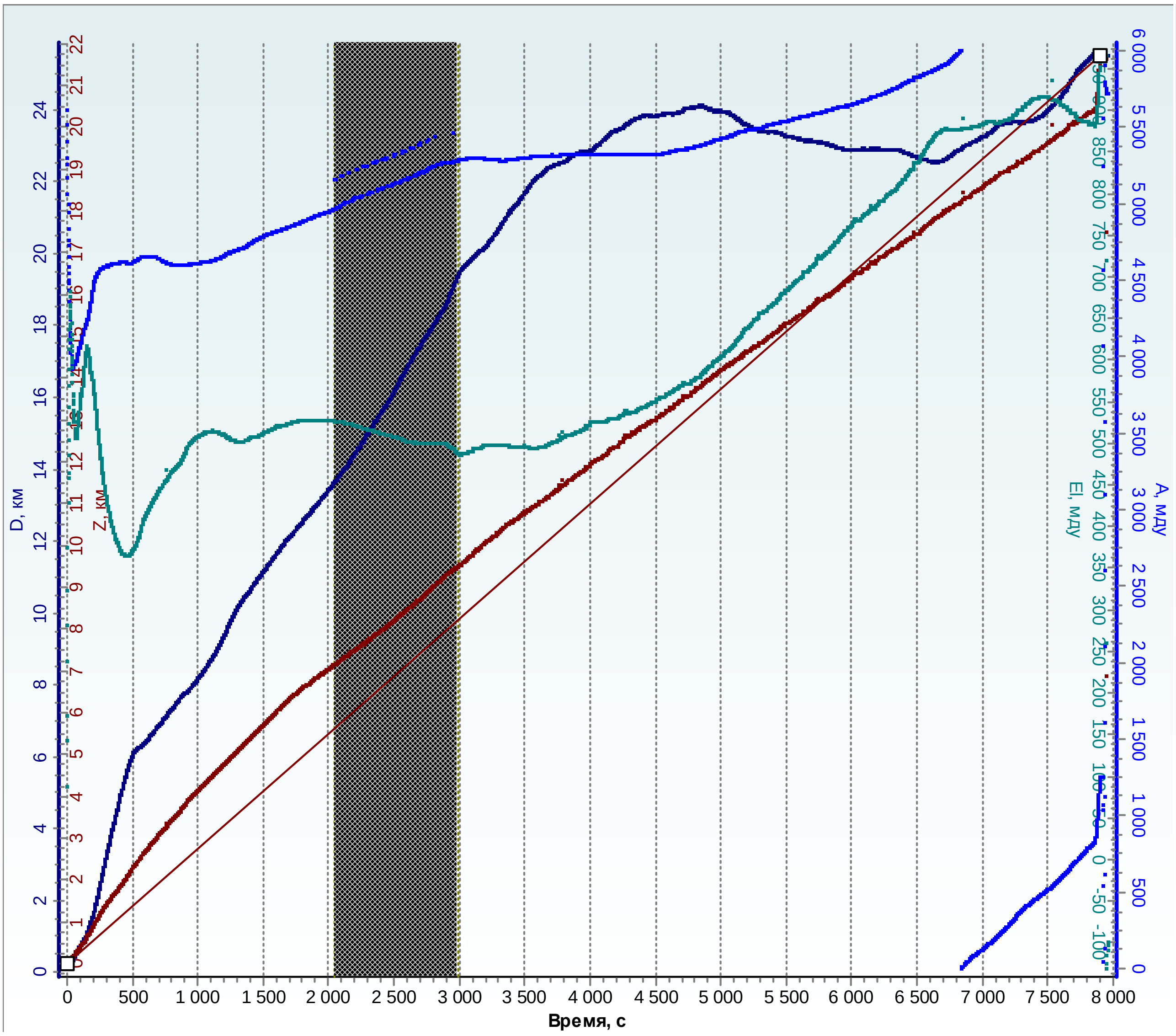

Figure 6.a. AVK radar tracking data (dark blue – slant range [km], brown – instantaneous height [km], blue – azimuth [1x0.06°] and cyan – elevation [1x0.06°]) – an improper adjustment of azimuth transducers.

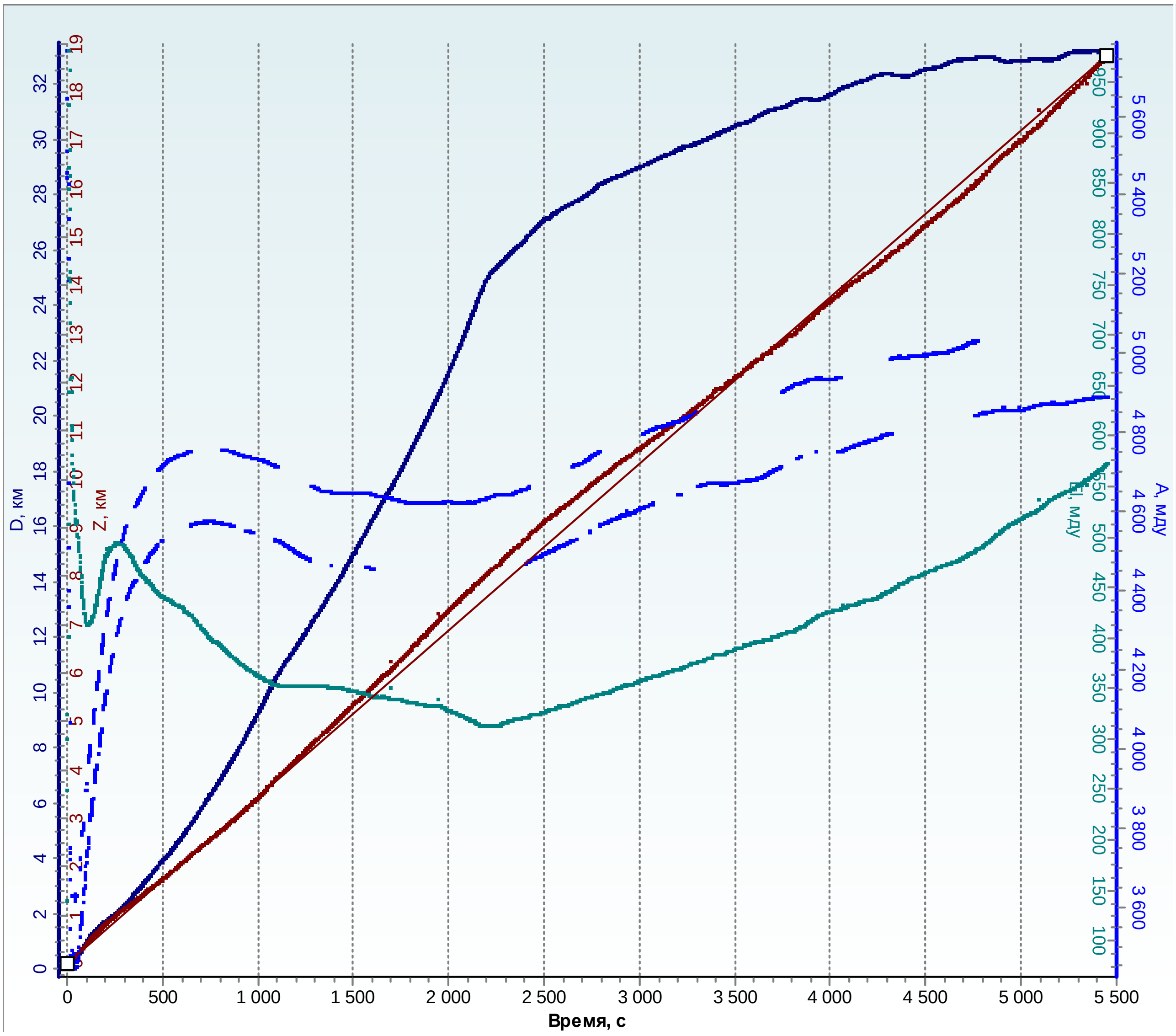

Figure 6.b. AVK radar tracking data (dark blue – slant range [km], brown – instantaneous height [km], blue – azimuth [1x0.06°] and cyan – elevation [1x0.06°]) – an improper adjustment of azimuth transducers , neglected case.

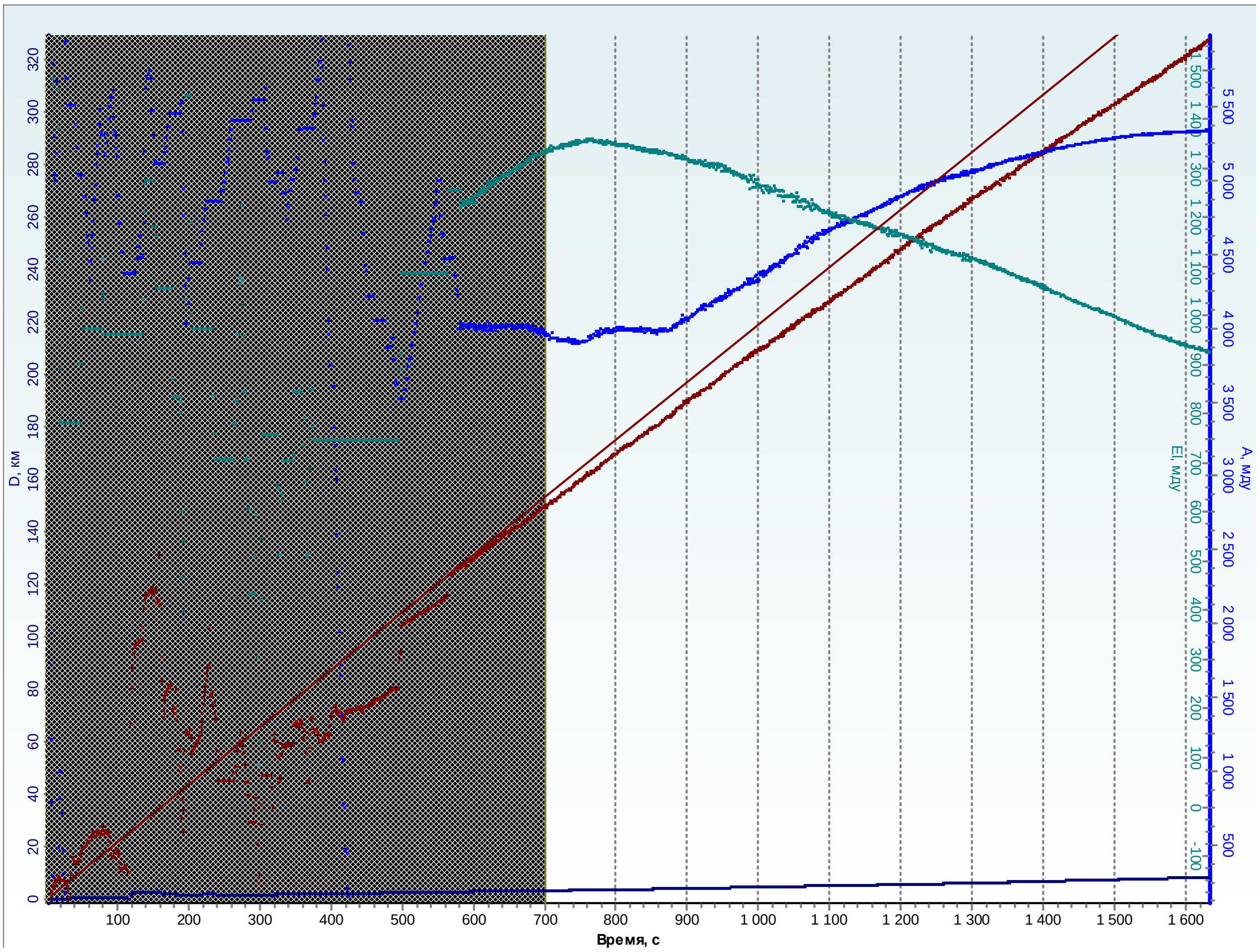

Figure 7.a. AVK radar tracking data (dark blue – slant range [km], brown – instantaneous height [km], blue – azimuth [1x0.06°] and cyan – elevation [1x0.06°]) – loss of a radiosonde just after release.

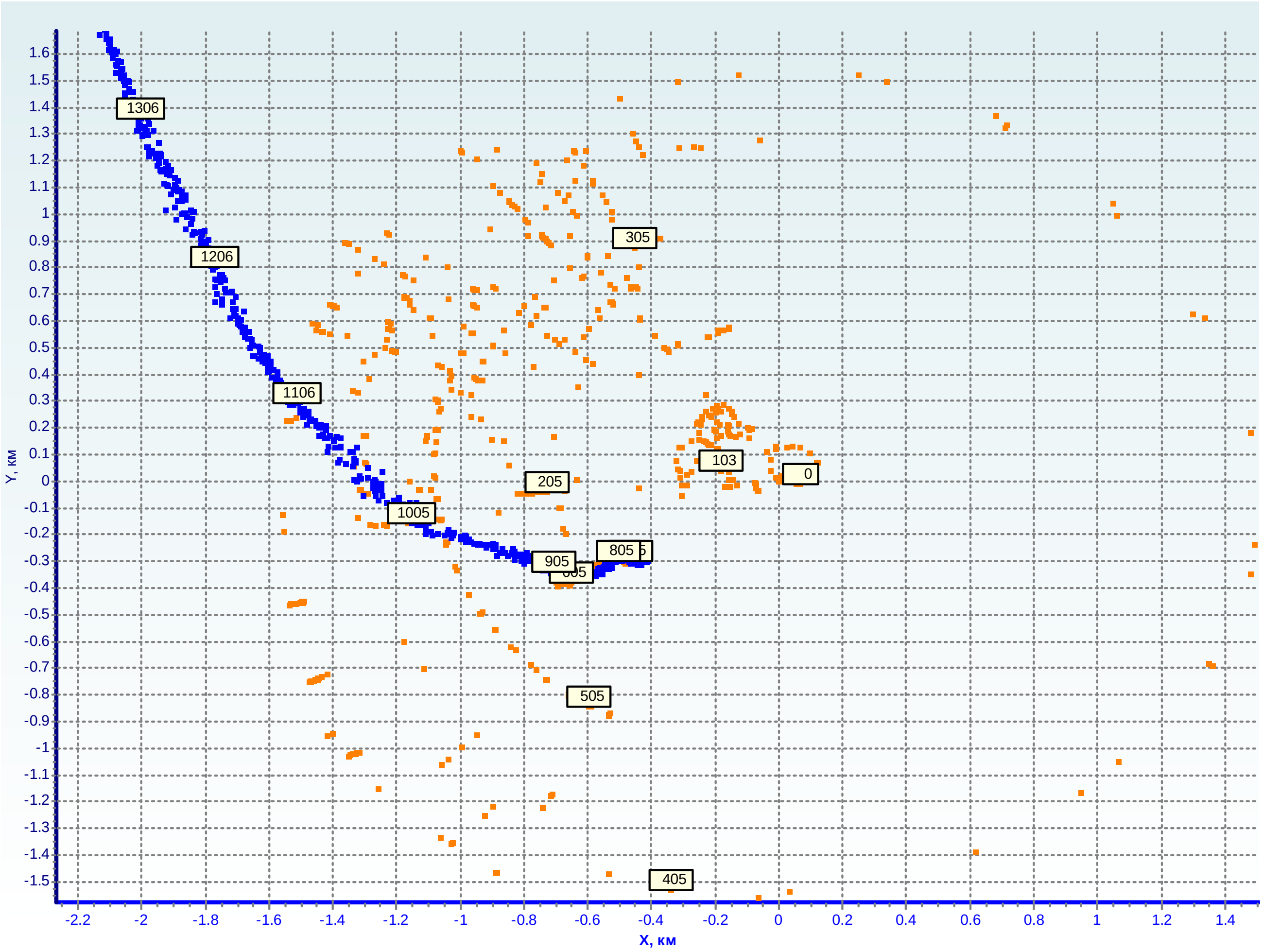

Figure 7.b. Same ascent as on Figure 7.a. AVK radar tracking data - horizontal projection [km], marks indicate elapsed ascent time [s] – loss of a radiosonde just after release. One can see how operator searched the radiosonde.

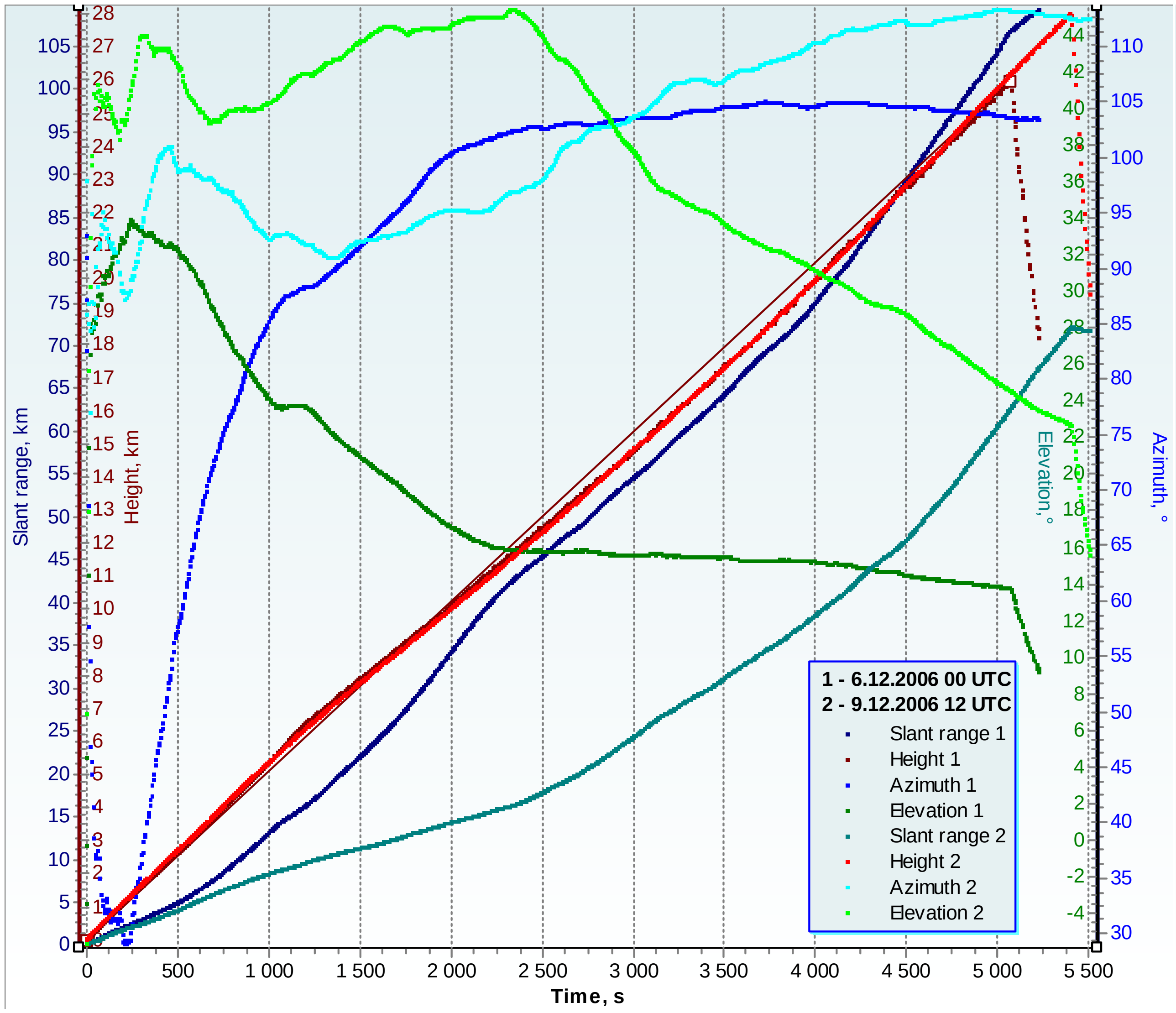

Figure 8. AVK radar tracking data – two successive ascents from the same station.

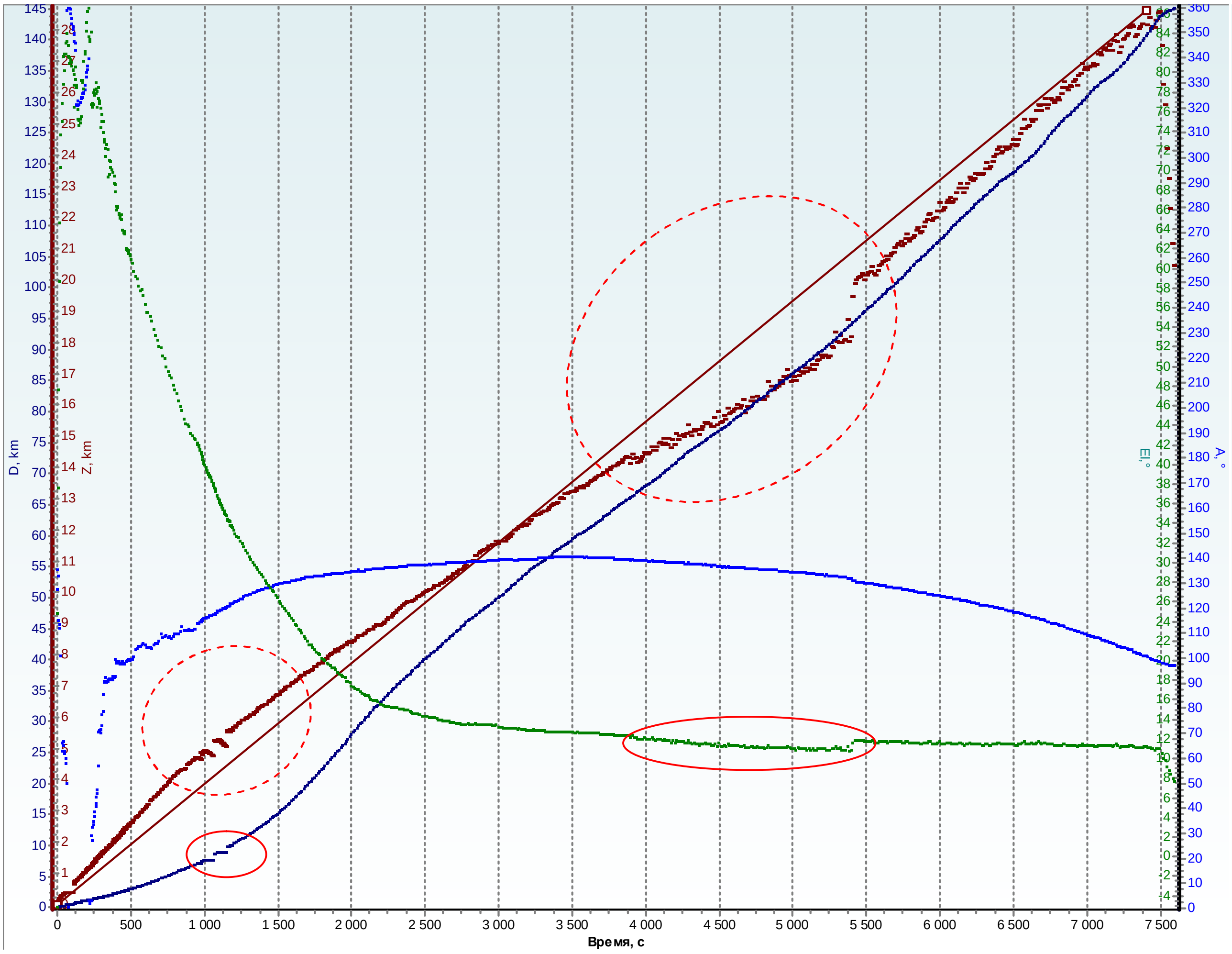

Figure 9.a. A radar tracking data (dark blue – slant range [km], brown – instantaneous height [km], blue – azimuth [°] and green – elevation [°])– peculiarities in height reflect tracking problems.

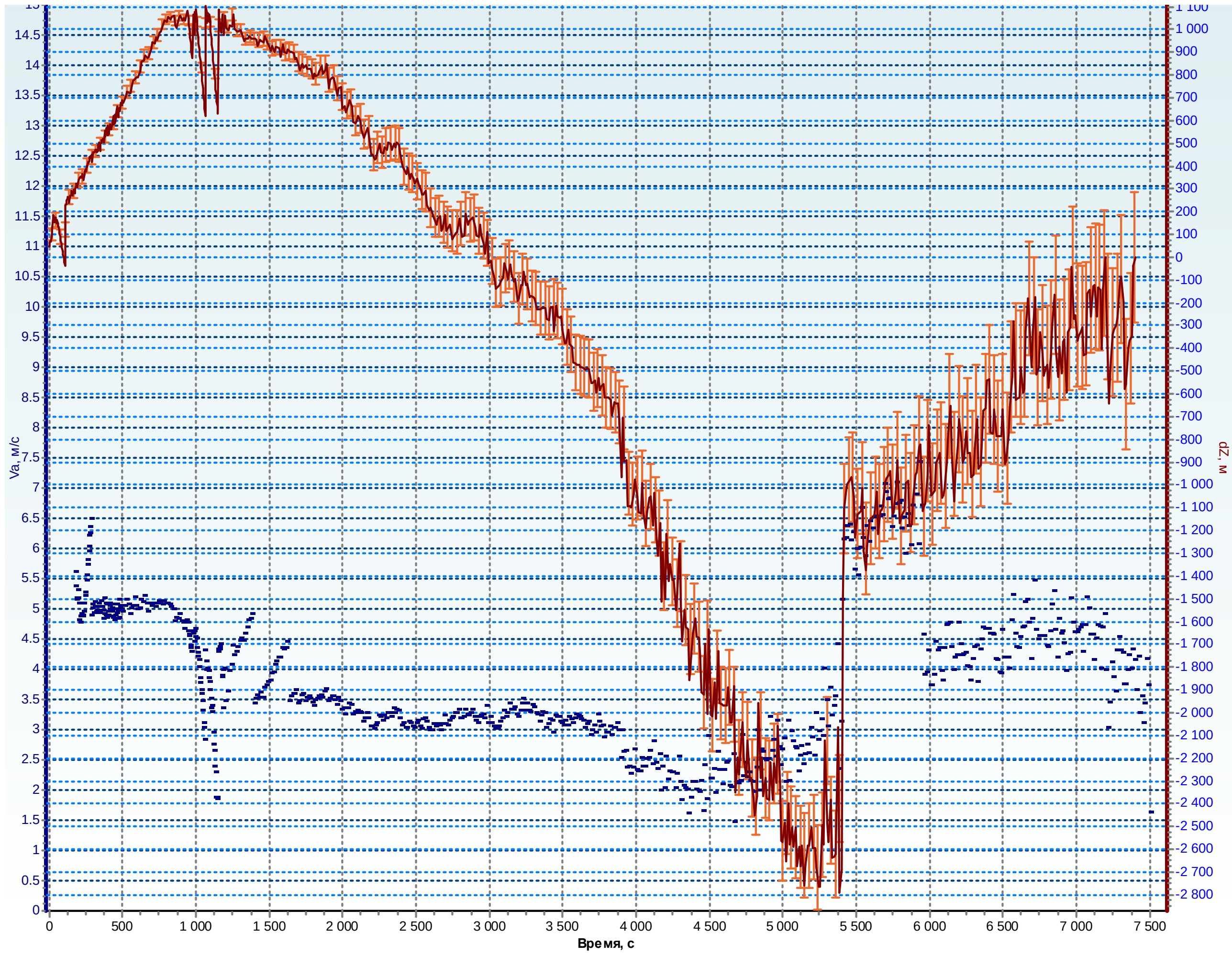

Figure 9.b. Same ascent as on Figure 9.a. Estimates of vertical ascent rate [m/s] (dark blue) and auxiliary quantity dZ [m] (brown). Vertical bar show estimates of standard deviation[9] of radar heights according to AVK nominal standard deviation of slant range and elevation measurements as high as 30 m and 0.12° respectively.

[9] As one can see – height error dramatically increases along with increase of slant range and decrease of elevation angle.

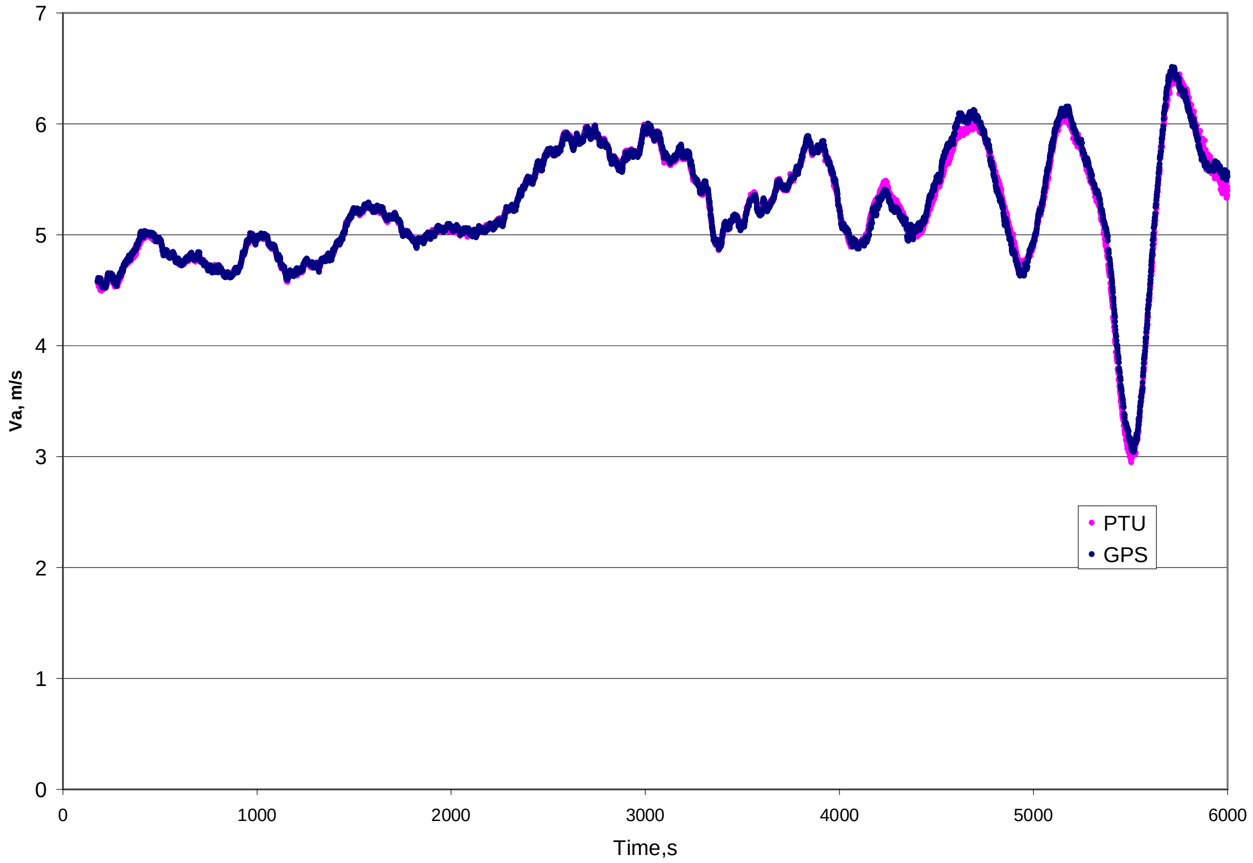

Figure 10.a. Estimates of vertical ascent rate [m/s] from Vaisala RS92 radiosonde flight raw data. PTU – from heights, derived from raw PTU with hydrostatic equation, GPS – from GPS heights.

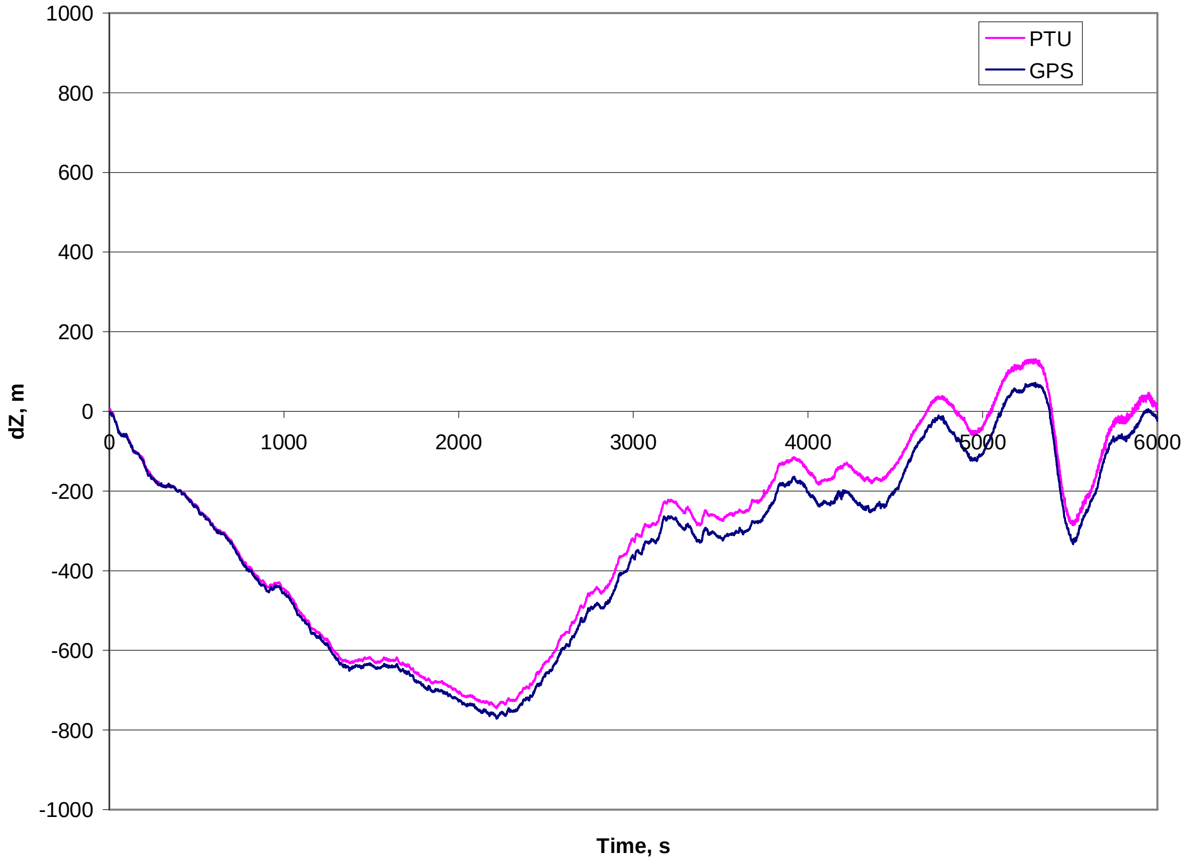

Figure 10.b. Same ascent as on Figure 10.a. Estimates of auxiliary quantity dZ [m] from Vaisala RS92 radiosonde flight raw data. PTU – from heights, derived from raw PTU with hydrostatic equation, GPS – from GPS heights.

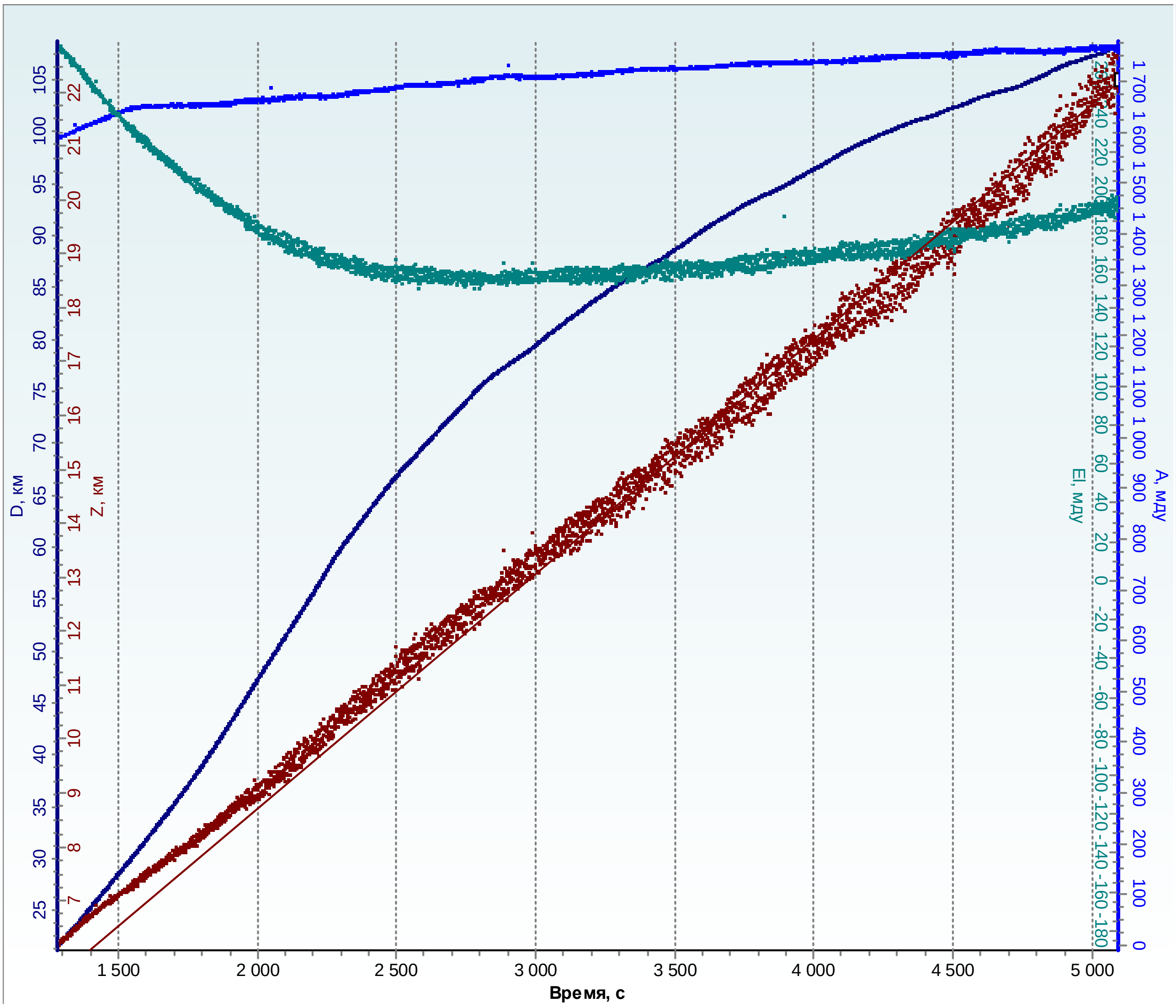

Figure 11.a. AVK radar tracking data (dark blue – slant range [km], brown – instantaneous height [km], blue – azimuth [1x0.06°] and green – elevation [°]) – height random errors increase as slant range increases and elevation angle decreases.

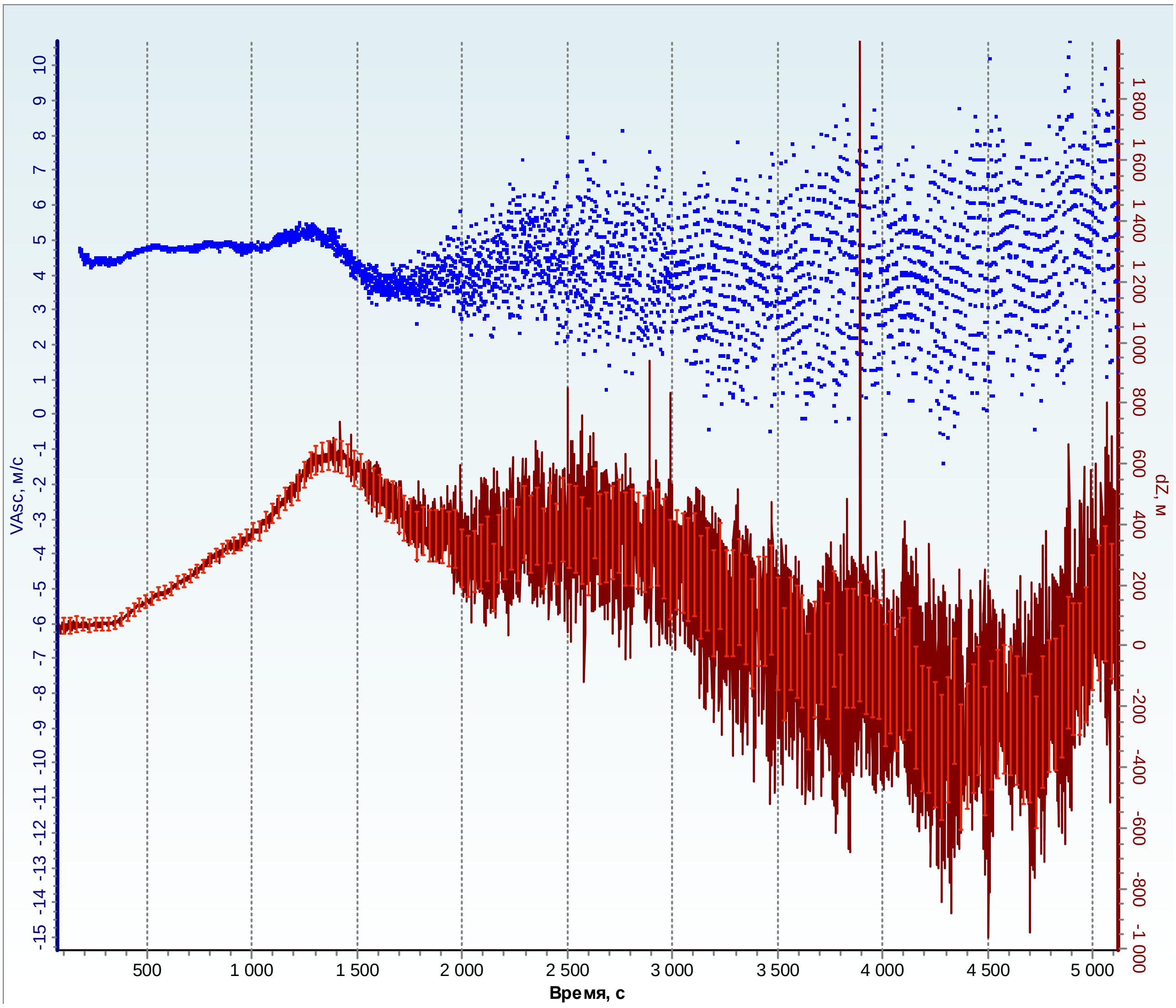

Figure 11.b. Same ascent as on Figure 11.a Estimates of vertical ascent rate [m/s] (dark blue) and auxiliary quantity dZ [m] (brown). Vertical bar show estimates of standard deviation of radar heights according to AVK nominal standard deviation of slant range and elevation measurements as high as 30 m and 0.12° respectively.

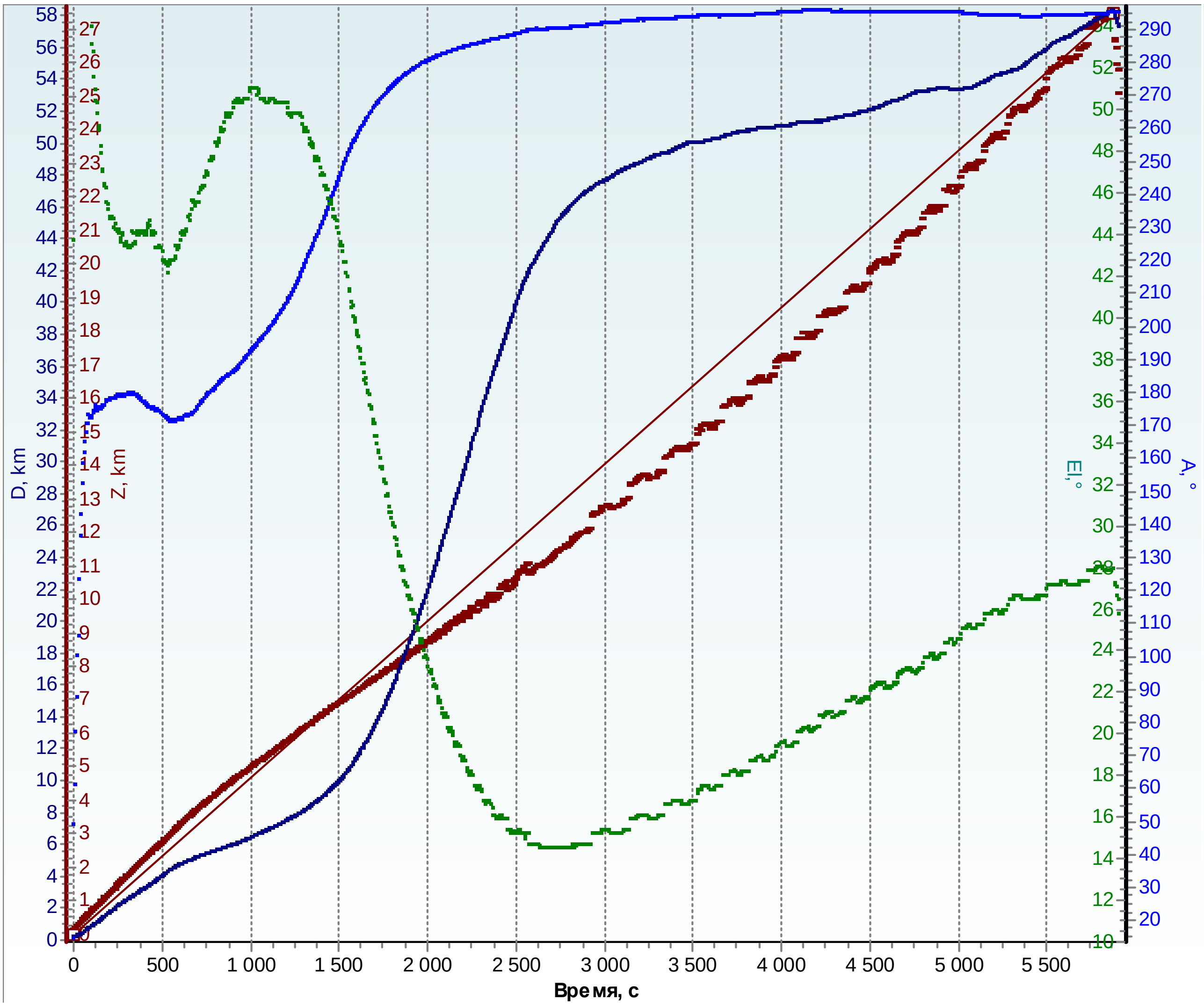

Figure 12.a. AVK radar tracking data (dark blue – slant range [km], brown – instantaneous height [km], blue – azimuth [1x0.06°] and green – elevation [°]) – malfunction of elevation-to-code digital transducer.

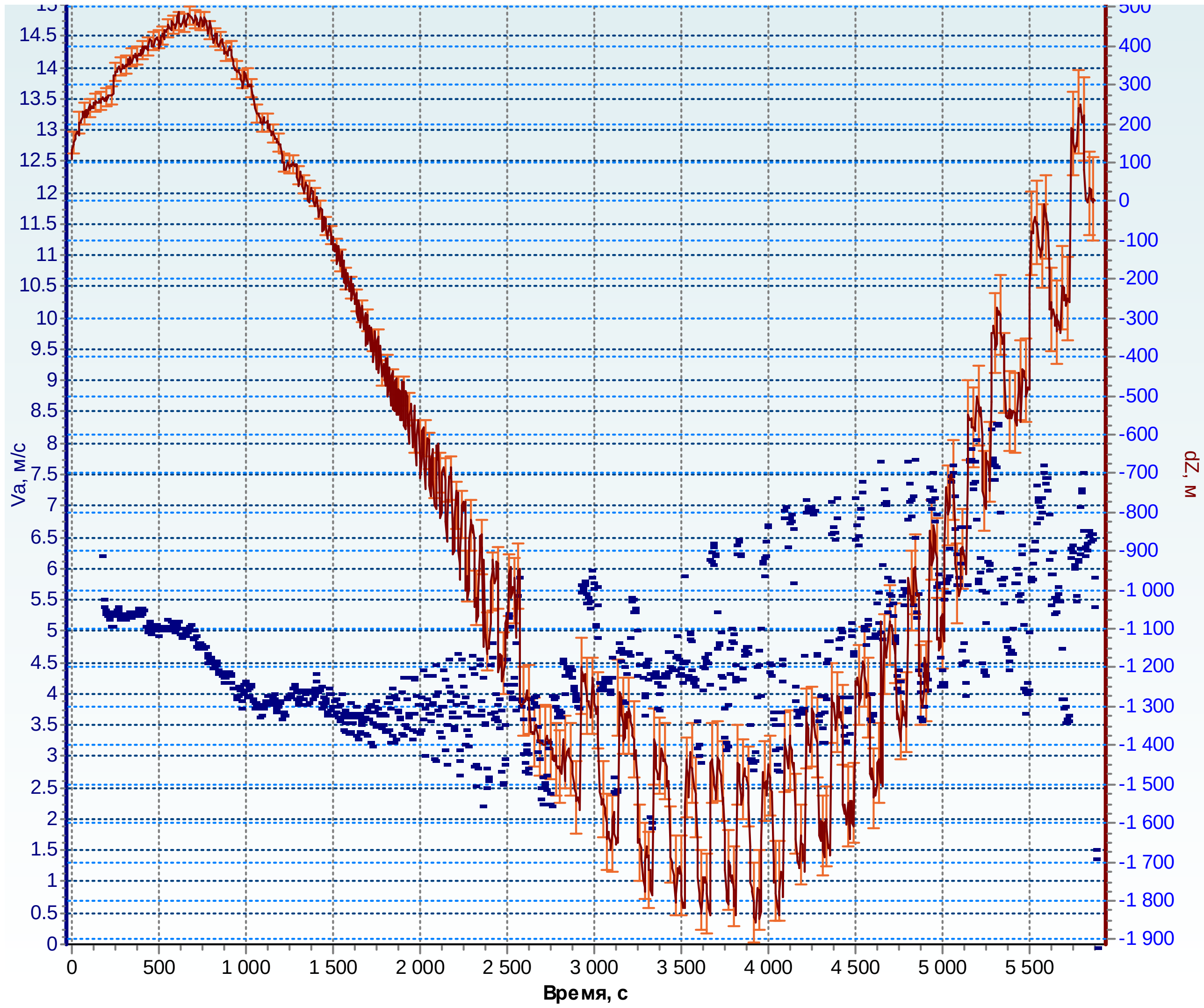

Figure 12.b. Same ascent as on Figure 12.a Estimates of vertical ascent rate [m/s] (dark blue) and auxiliary quantity dZ [m] (brown). Vertical bar show estimates of standard deviation of radar heights according to AVK nominal standard deviation of slant range and elevation measurements as high as 30 m and 0.12° respectively.

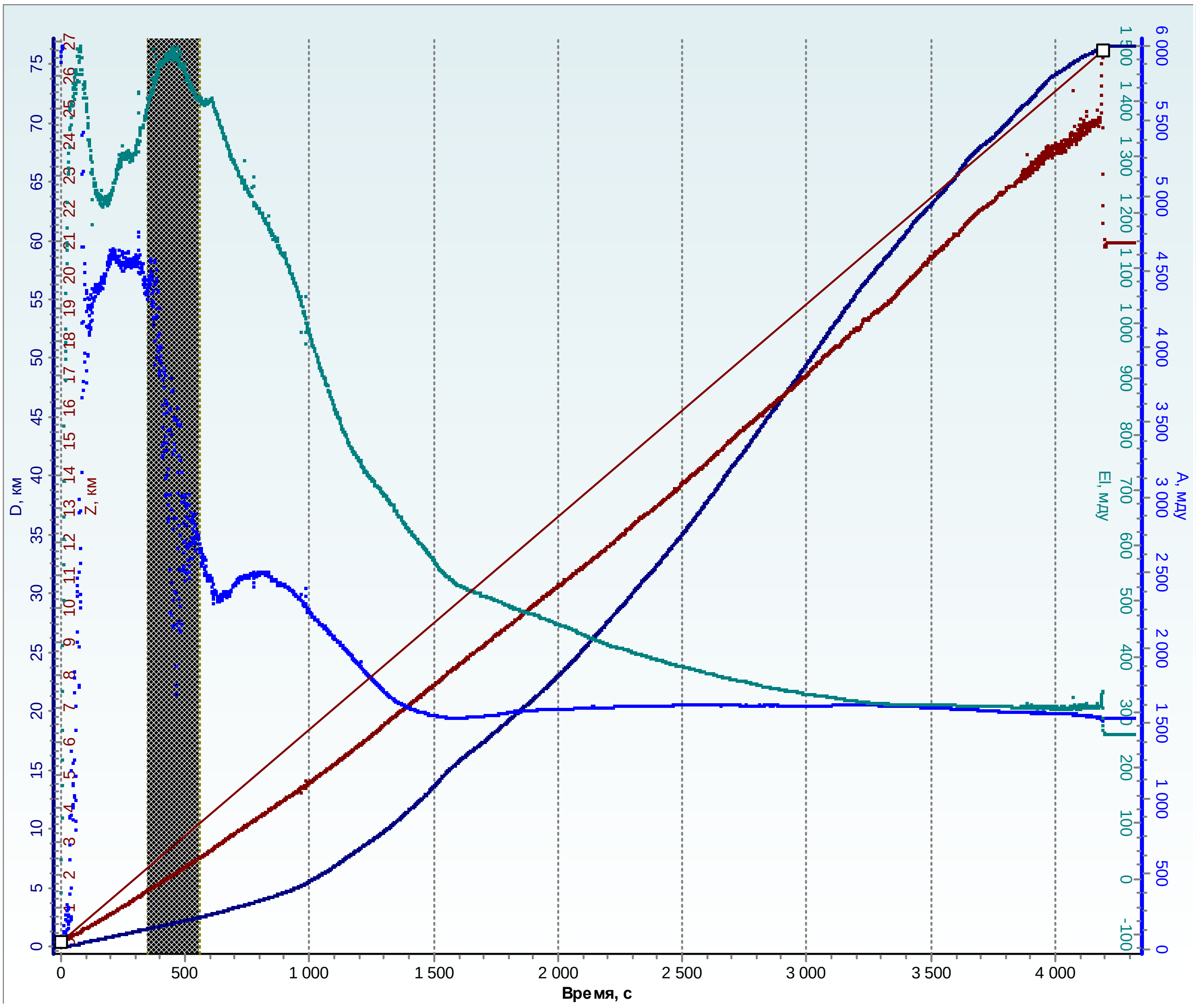

Figure 13.a. AVK radar tracking data (dark blue – slant range [km], brown – instantaneous height [km], blue – azimuth [1x0.06°] and cyan – elevation [1x0.06°]) – poor azimuthal tracking.

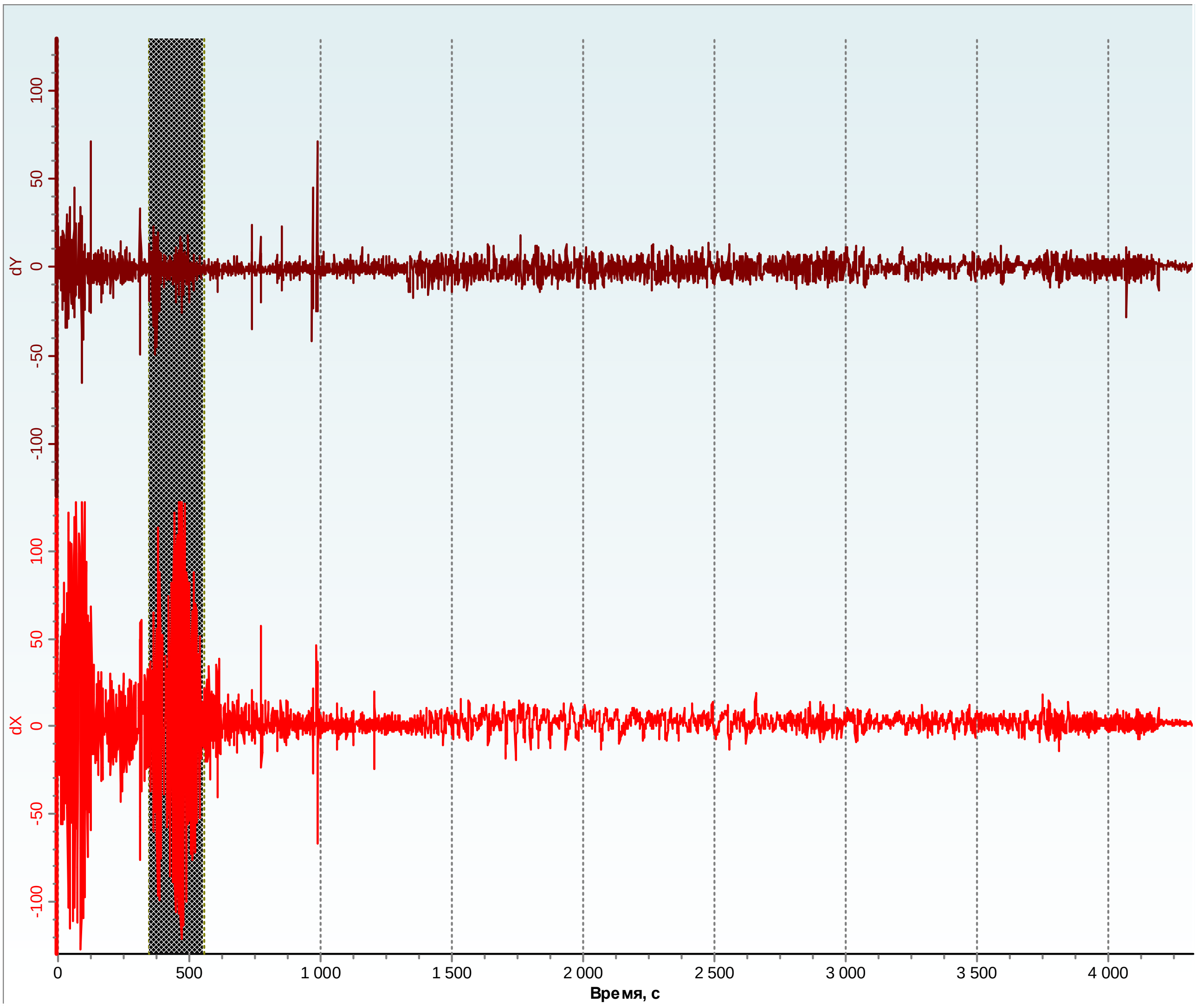

Figure 13.b. Same ascent as on Figure 13.a. AVK radar tracking data (brown – elevation tracking error, red – azimuth tracking error, both are nondimensional) – poor azimuthal tracking.

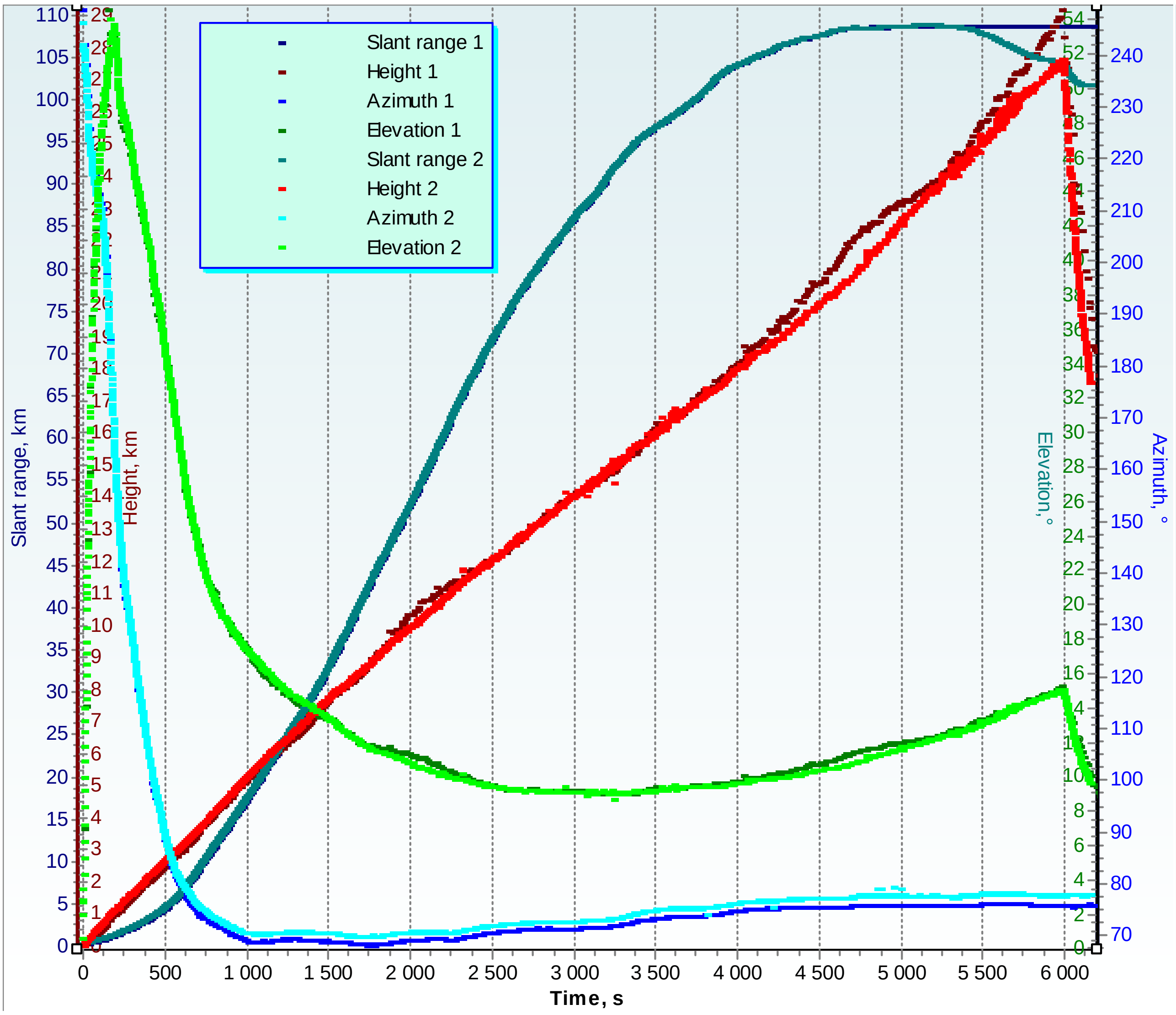

Figure 14. Two radars tracking data – a twin flight. The second radar provides more consistent data.

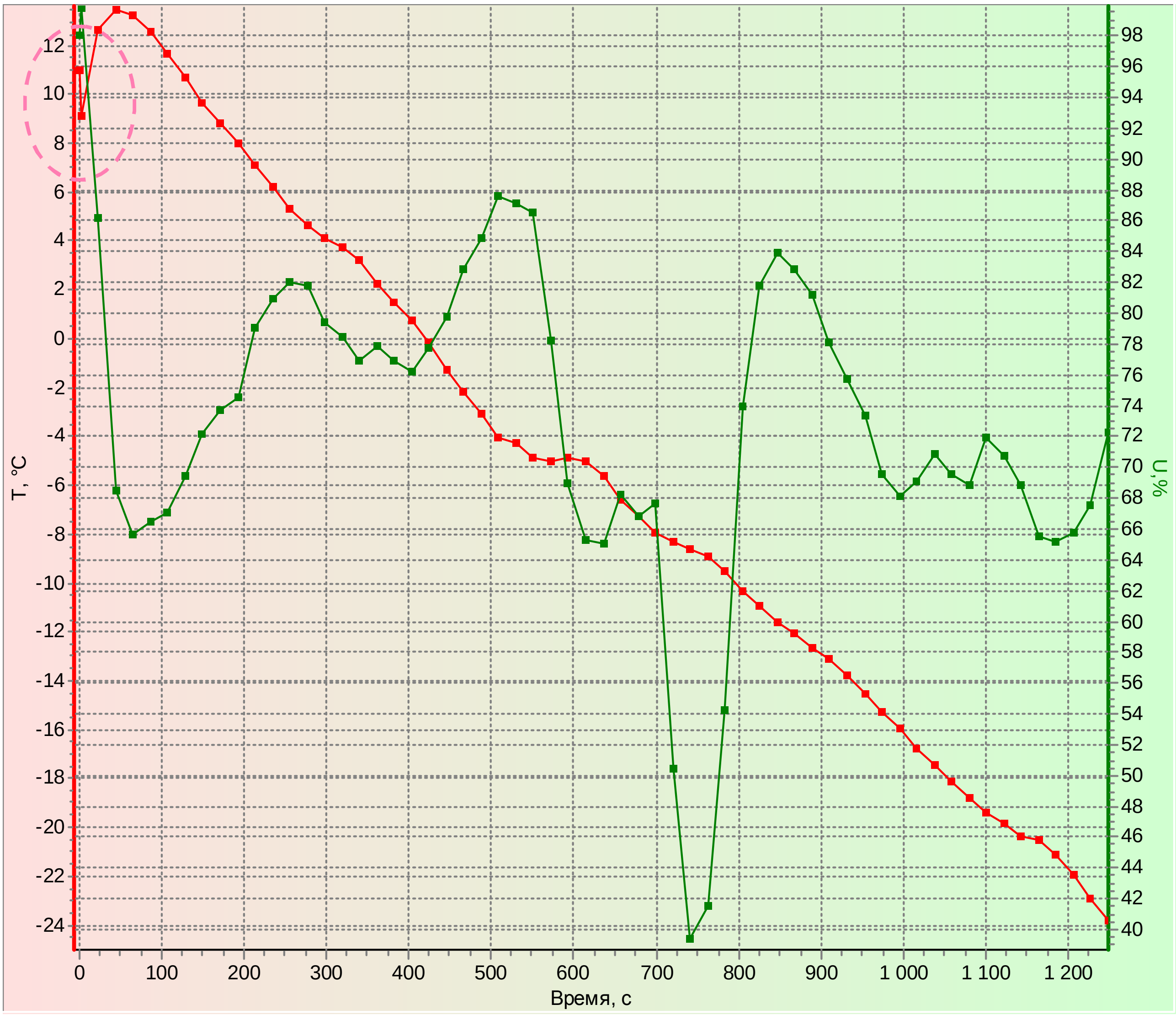

Figure 15. Primary temperature [°C] (red) and humidity [%RH] (green), calculated directly from telemetry frequencies and surface temperature and humidity (at 0 s). Surface temperature is 11 °C and the first temperature reading at 3 s is 9 °C.

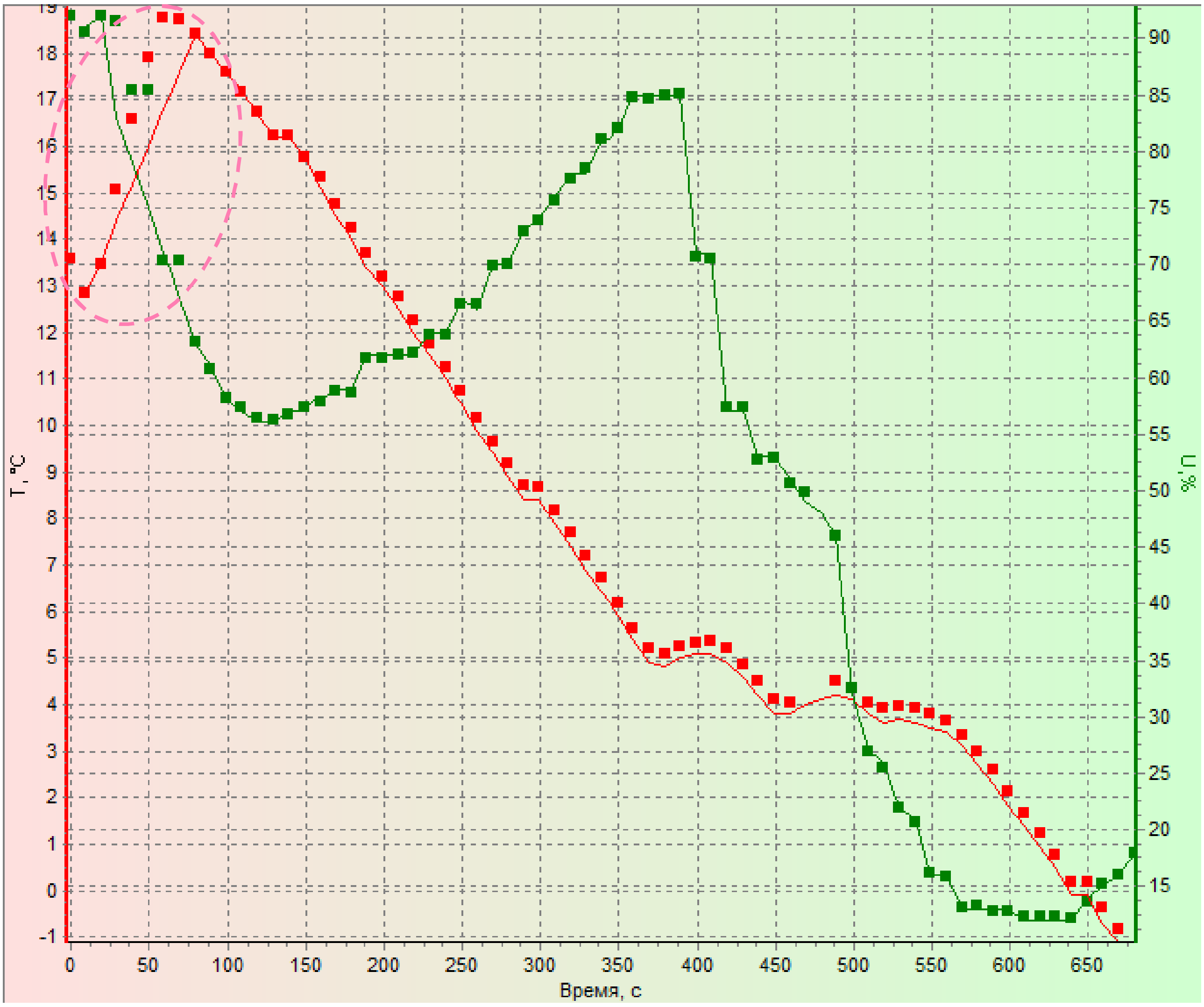

Figure 16. Primary (points) and edited (lines) temperature [°C] and humidity [%RH] – strong inversion was smoothed.

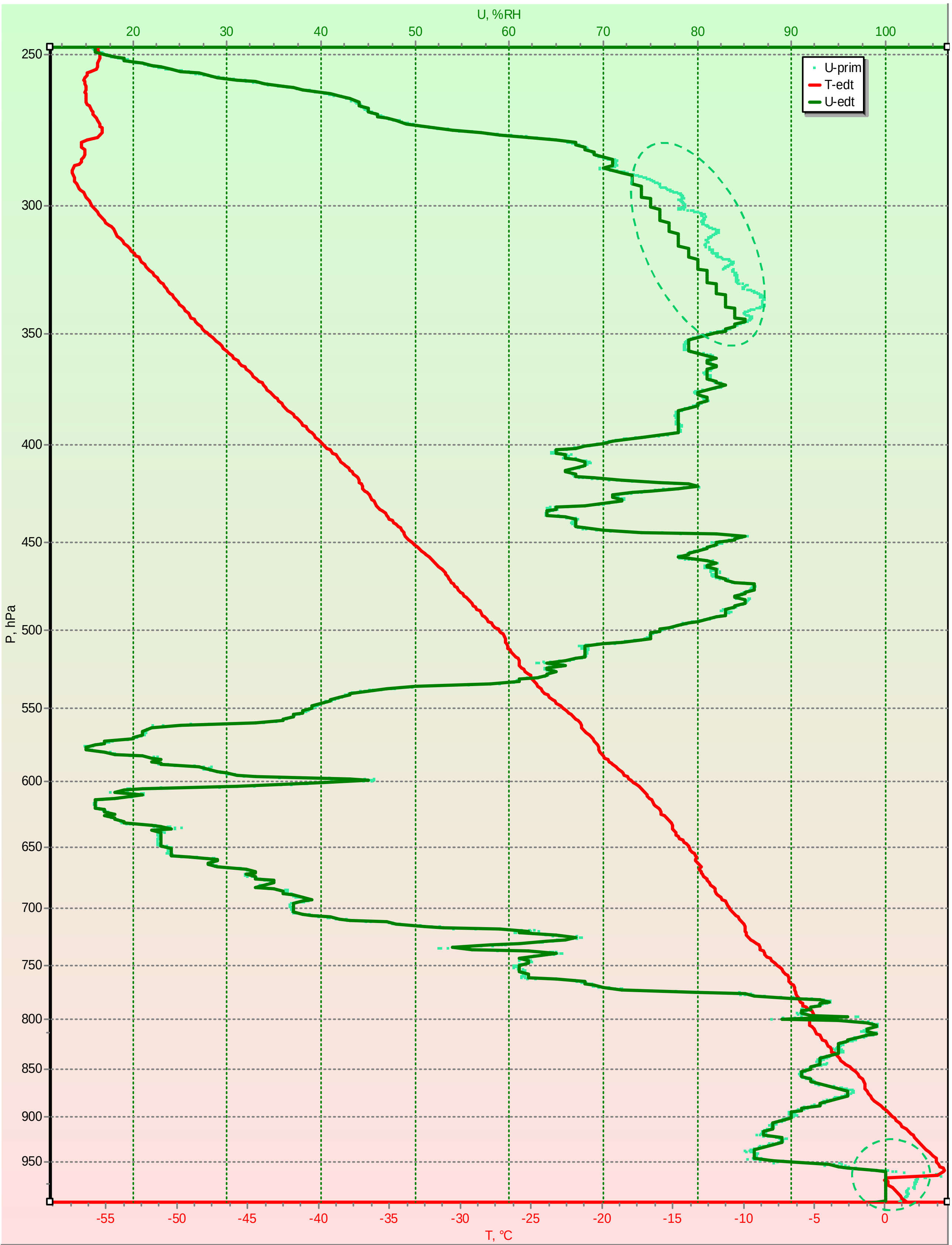

Figure 17. Primary (U-prim) and edited (U-edt) humidity [%RH] and edited temperature [°C] (T-edt) – humidity values with super saturation over ice and over water were truncated.

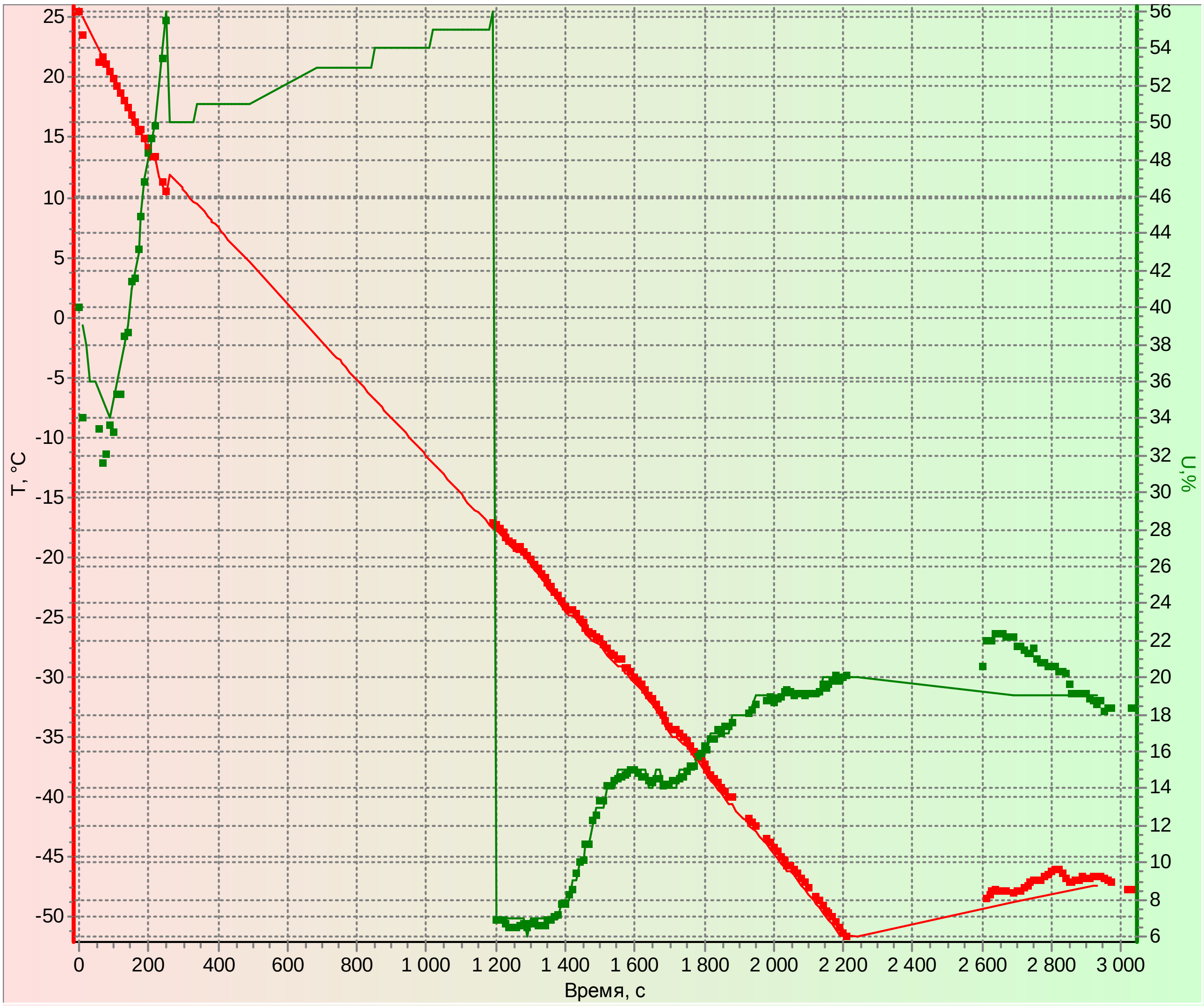

Figure 18. Primary (points) and edited (lines) temperature [°C] and humidity [%RH] – lost data filling.

Figure 19. Primary (points) and edited[10] (lines) temperature [°C] and humidity [%RH] – massive data loss and outliers.

[10] With radiation correction applied

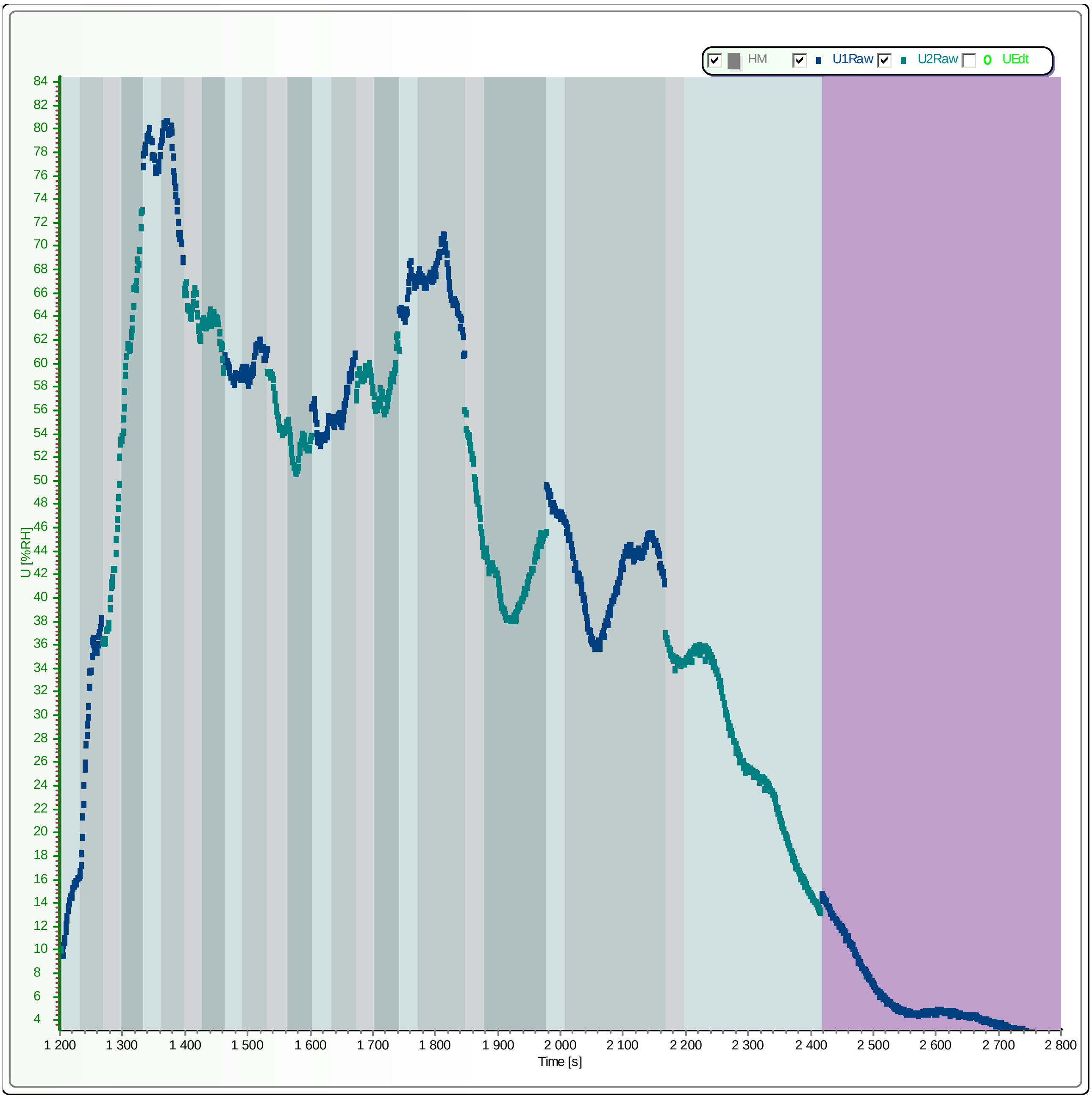

Figure 20. Alternative humidity measurements by two RS92 H-HUMICAP humidity sensors. Light and dark bands correspond to heating and cooling of an alternative sensor.